Longitudinal Study of External Validity of the PARCC Performance Levels:

Phase 1 Report

Authors:
Jeffrey Steedle, Sarah Quesen,
and Aimee Boyd
Pearson
March 1, 2017

Table of Contents

Executive Summary

On the Partnership for Assessment of Readiness for College and Careers (PARCC) assessments, the attainment of performance level 4 is intended to indicate college readiness or being "on track" to college and career readiness. Students who achieve Level 4 are should have a .75 probability of attaining at least a C in entry-level, credit-bearing, post-secondary coursework. This report describes the first phase of a two-part longitudinal external validity study of claims about the alignment of PARCC Level 4 to college readiness.

PARCC assessment scores from the 2014–2015 and 2015–2016 academic years were used to examine associations between PARCC performance levels and college readiness benchmarks established by the College Board and ACT. Parametric (OLS) and nonparametric (LOESS) regression estimates measured the relationship between the PARCC and external test scores. The PARCC Level 4 benchmark was used to estimate the expected score on an external test, and vice versa. Assessment scores were dichotomized for additional analyses, with 0 indicating performance below the benchmark and 1 indicating performance at or above the benchmark. Cross-tabulation tables provided classification agreement among tests. Logistic regression modeled the relationship between students' PARCC scores and their probabilities of meeting the external assessment benchmark, and vice versa.

These methods were used to make the following comparisons in mathematics: Algebra 1 and PSAT10 Math; Geometry and PSAT10 Math; Algebra II and PSAT10 Math; Algebra II and PSAT/NMSQT Math; Algebra II and SAT Math; and Algebra II and ACT Math. The classification agreement (meeting the benchmark on both tests or not meeting the benchmark on both tests) ranged from 62.5% to 86.5%. **The overall trend indicated that students who met the benchmark on a PARCC mathematics assessment were likely to meet or exceed the benchmark on an external test (probabilities ranged from .509 to .886). However, students who met the benchmark on the external test had relatively low probabilities of meeting the PARCC mathematics benchmark (.097 to .310).**

The following comparisons were made in English language arts/literacy: grade 9 and PSAT10 evidence-based reading and writing (EBRW); grade 10 and PSAT10 EBRW; grade 10 and PSAT/NMSQT EBRW; grade 10 and SAT EBRW; grade 11 and PSAT/NMSQT EBRW; grade 11 and SAT EBRW; grade 11 and ACT English; and grade 11 and ACT reading. In the majority of comparisons, the trend in ELA/L results was similar to mathematics. **The classification agreements ranged from 67.3% to 79.7%. Students meeting the PARCC ELA/L benchmark had probabilities between .667 and .825 of meeting the benchmark on the external assessment.**

However, a student taking the external test had lower probabilities of meeting the benchmark on the PARCC ELA/L assessments (.326 to .513)

Overall, results indicated that a student meeting the benchmark on the PARCC test had a high probability of making the benchmark on the external test, but the converse did not hold for students meeting the benchmark on the external test, for the majority of comparisons (Table 0.1). These results suggests that meeting the PARCC benchmark is an indicator of academic readiness for college. However, it may be that students who meet the PARCC benchmark have a greater than .75 probability of earning a C or higher in first-year college courses.

Phase 1 is a preliminary study using indirect comparisons; therefore, there are limitations to interpretations. There may not have been sufficient overlap in content between tests, the samples may not have reflected the population of PARCC examinees, and the timing of assessments may not have provided a level playing field for these comparisons. A future analysis could assess the sensitivity of the results to the sample by creating samples from the available data with demographic and ability distributions similar to those of the target PARCC populations.

Phase 2 of this study (to occur in 2018) will use longitudinal data including academic performance in entry-level college courses for students who took PARCC assessments during high school. This second phase may lend additional support for the validity of the PARCC Level 4 performance level as a predictor of postsecondary academic success.

Table 0.1

Results Summary

Result	Comparison
PARCC benchmark is more rigorous	Geometry and PSAT10 Math Algebra II and PSAT10 Math Algebra II and PSAT/NMSQT Math Algebra II and SAT Math Grade 10 ELA/L and PSAT10 EBRW Grade 10 ELA/L and PSAT/NMSQT EBRW Grade 10 ELA/L and SAT EBRW Grade 11 ELA/L and PSAT/NMSQT EBRW Grade 11 ELA/L and SAT EBRW
Benchmarks are close or PARCC is slightly more rigorous	Algebra II and ACT Math Grade 9 ELA/L and PSAT10 EBRW
Benchmarks are close or PARCC is slightly less rigorous	Algebra I and PSAT10 Math
External benchmark is more rigorous	Grade 11 ELA/L and ACT English Grade 11 ELA/L and ACT Reading

1. Introduction

The performance level setting (PLS) process was designed to help students, parents, teachers and policymakers understand how students performed relative to the content standards for college and career readiness. During the summer of 2015, educators and content area experts gathered to determine the threshold scores that would define the boundaries between five performance levels on the Partnership for Assessment of Readiness for College and Careers (PARCC) assessments. It was intended that performance at or above the Level 4 threshold would indicate meeting expectations reflected by the content standards and being "on track" for the next grade level and, ultimately, for college and careers. The extent to which stakeholders can make valid interpretations of performance level designations depends on how accurately the thresholds distinguish between the performance levels as defined by the PARCC performance level descriptors,[1] which explain the knowledge, skills, and abilities demonstrated by students at each level.

For most PARCC assessments, meeting the Level 4 threshold score is intended to indicate readiness for the next level of coursework and being "on track" to college and career readiness. In the case of capstone PARCC assessments (grade 11 ELA/Literacy, Algebra II, and Integrated Mathematics III), meeting the Level 4 threshold scores is intended to indicate academic readiness for postsecondary coursework in the associated content area. In addition, meeting the Level 4 threshold on capstone assessments should be predictive of postsecondary academic success. This *PARCC College- and Career-Ready Determination Policy in English Language Arts/Literacy and Mathematics* (a*dopted by the PARCC Governing Board and Advisory Committee on College Readiness October 25, 2012*)[2] statement forms the basis for validation studies of the efficacy of PARCC's College- and Career-Ready Determinations:

> Students who earn a College- and Career-Ready Determination by performing at Level 4 in grade 11 ELA/literacy and enroll in College English Composition, Literature, and technical courses requiring college-level reading and writing have approximately a 0.75 probability of earning college credit by attaining at least a grade of C or its equivalent in those courses.

> Students who earn a PARCC College- and Career-Ready Determination by performing at Level 4 in Algebra II or Mathematics III and enroll in College Algebra, Introductory College Statistics, and technical courses requiring an equivalent level of mathematics

[1] http://www.parcconline.org/assessments/test-design/mathematics/math-performance-level-descriptors
http://www.parcconline.org/assessments/test-design/ela-literacy/ela-performance-level-descriptors
[2] http://www.parcconline.org/files/79/College%20and%20Career%20Ready/92/PARCCCCRDPolicyandPLDsFINAL.pdf

have approximately a 0.75 probability of earning college credit by attaining at least a grade of C or its equivalent in those courses.

A 0.75 probability of earning a C or higher was chosen for two reasons: a C is the minimum grade required to receive course credit in most postsecondary institutions, and 0.75 is a reasonably high standard that is comparable to criteria used in the development of other college readiness benchmarks. If most students who meet or exceed the Level 4 threshold eventually demonstrate postsecondary academic success (and most students who do not meet the Level 4 threshold score do not demonstrate similar success), this would provide evidence supporting the validity of the Level 4 threshold as an indicator of academic readiness for college.

The PARCC high school ELA/L grade 11 and Algebra II assessments were evaluated on test content based on criteria developed by the Council of Chief State School Officers (CCSSO) for assessments designed to measure college- and career-ready content standards (CCSSO, 2014). A study by Human Resources Research Organization (HumRRO) evaluated the extent to which the PARCC assessments match the CCSSO criteria, *Criteria for Procuring and Evaluating High Quality Assessments* (Schultz, Michaels, Dvorak, & Wiley, 2016). The study assigned two ratings, Content and Depth, to each PARCC assessment. The study found both ELA/L grade 11 and Algebra II to be well aligned with college and career readiness standards.

The purpose of this research is to evaluate the validity of PARCC Level 4 as an indicator of being college ready or "on track" to being college ready. The research is being conducted in two phases. In this first phase, PARCC high school assessment data from the 2014–2015 and 2015–2016 academic years were used to examine the associations between the PARCC performance levels and college readiness benchmarks established by the College Board (PSAT and SAT)[3] and ACT (the ACT assessment).[4] This phase of the study might be considered "indirectly longitudinal" since the College Board and ACT benchmarks are based on longitudinal data. The second phase will occur in 2018, at which point robust longitudinal data will be available. Those data will include academic performance in entry-level college courses for students who took PARCC assessments during high school. With those data, performance on the PARCC assessments and corresponding performance level assignments can be associated with postsecondary academic success. A strong association would support the validity of interpreting PARCC assessment results as predictors of postsecondary academic success.

---

[3] https://collegereadiness.collegeboard.org/about/scores/benchmarks
[4] https://www.act.org/content/act/en/education-and-career-planning/college-and-career-readiness-standards/benchmarks.html

This document describes the outcomes of the first phase of the study for investigating the validity of the PARCC performance levels as indicators of academic readiness for college.

2. Method

Study Data

In this study, PARCC Level 4 was compared to college readiness benchmarks on external assessments. There are straightforward connections between PARCC Level 4 on capstone assessments (Grade 11 ELA/L and Algebra II) and the SAT and ACT benchmarks because they are all intended to indicate college readiness for students approaching high school graduation. However, a study examining only these connections is limited by available data. Namely, not all PARCC states administered capstone assessments in 2014–2015. In states that did, many students who took Algebra II were tenth graders, and such students would not have taken the SAT or ACT by the time of this study in 2016.

To address these limitations, this study also examined connections between Level 4 on PARCC assessments administered earlier (e.g., at grades 8–10) and external assessments with benchmarks indicating whether students are "on track" to college readiness. PSAT and ACT Aspire[5] readiness benchmarks are vertically aligned to the SAT and ACT benchmarks, respectively. Results suggesting that PARCC "on track" benchmarks are consistent with other "on track" benchmarks supports the notion that attaining PARCC Level 4 on Algebra I, Geometry, Grade 9 ELA/L, or Grade 10 ELA/L provides a valid indication that students are on track to being college ready. Such findings may also indirectly validate the PARCC college readiness benchmarks for capstone assessments because of the vertical articulation that occurred for PARCC performance levels during the performance level setting process. Note that comparisons of performance on different assessments depend on the extent to which those assessments measure mastery of similar content. Although PARCC, College Board, and ACT assessments do not measure the same body of content, convergent performance (e.g., meeting college readiness benchmarks on both PARCC and the SAT) supports the validity of PARCC's college readiness benchmarks.

College Readiness Benchmarks
College Board Benchmarks
Administration of the redesigned SAT and PSAT began in the 2015–2016 school year (first PSAT administration in October 2015, first SAT administration in March 2016). Compared to the

---

[5] https://collegereadiness.collegeboard.org/about/scores/benchmarks
https://www.discoveractaspire.org/assessments/standards-benchmarks/

older test design, writing scores are no longer incorporated into total scores, and scores from the PSAT and SAT are reported on the same scale. The tests include two sections: Math and Evidence-Based Reading and Writing. The redesigned SAT benchmarks indicate a 0.75 probability of "achieving at least a C in a set of first-year, credit-bearing college courses" (College Board, 2015, p. 6). Whereas, prior SAT benchmarks were linked to overall first-year grade-point averages, scores on the revised SAT will be linked to performance in specific, first-year college courses. The PSAT benchmarks reflect the SAT benchmarks adjusted for average rates of improvement from year to year (College Board, 2015). The PSAT and SAT benchmarks are shown in Table 2.1. Note that the PARCC and revised SAT benchmarks are similarly defined (0.75 probability of earning a C or higher).

College Board published results from a pilot predictive validity study for the redesigned SAT (Shaw, Marini, Bear, Shmueli, Young, & Ng, 2016). In that study, the redesigned SAT was administered to entering first-year college students in fall 2014. Results revealed that the redesigned SAT correlated with college grades in a manner similar to previous versions of the SAT. A full validity study using new longitudinal data sets will not be published until 2019. Data gathered for this study included a mix of older and redesigned SAT and PSAT results. Results from the older tests were transformed to the new scale using concordance tables provided by the College Board.

Table 2.1

Revised PSAT and SAT College Readiness Benchmarks (2015–2016 school year and beyond)[6]

PARCC Performance Level	PSAT Benchmark*	SAT Benchmark
Level 4 in Mathematics	510 on Math (Grade 11) 480 on Math (Grade 10)	530 on Math
Level 4 in ELA/Literacy	460 on Evidence-Based Reading and Writing (Grade 11) 430 on Evidence-Based Reading and Writing (Grade 10)	480 on Evidence-Based Reading and Writing

* PSAT 10 is administered to tenth graders and PSAT/NMSQT is administered to eleventh graders.

ACT Benchmarks

The ACT college readiness benchmarks (Table 2.2) indicate a 0.50 probability earning a B or higher and a 0.75 probability of earning a C or higher in specific, first-year college courses (Allen, 2013). The PARCC and ACT benchmarks are similarly defined (0.75 probability of earning a C or higher). Note that data from ACT Aspire were requested, but none were received.

---

[6] https://collegereadiness.collegeboard.org/about/scores/benchmarks

Table 2.2
ACT College Readiness Benchmarks[7]

PARCC Performance Level	ACT Benchmark
Level 4 in Mathematics	22 on ACT Mathematics (benchmark for College Algebra)
Level 4 in ELA/Literacy	18 on ACT English (benchmark for English Composition) 22 on ACT Reading (benchmark for Social Studies)

Data Preparation

For this study, data from the following external assessments were received: SAT (class of 2016), ACT (class of 2016), PSAT (2014–2015), and PSAT (2015–2016). Five states provided external assessment data files. Maryland, New Jersey, and New Mexico provided data files for all four assessments. Massachusetts provided SAT (class of 2016) data, and Rhode Island provided PSAT data from both requested years. To avoid the transmission of personally identifiable information, each participating state stripped those data from the files and added the PARCC student identifier. Possible reasons for not providing data included lack of data from College Board or ACT, lack of permission to share the data for this study, or not administering the high school PARCC assessments.

There have been recent revisions to the PSAT and SAT assessments, and data from the older and revised tests were both provided for this study. To combine these data, it was necessary to transform older PSAT and SAT scores to the revised score scales. For ELA, older PSAT Critical Reading and PSAT Writing scores were added together and a concordance table was applied to put the combined scores on the revised PSAT Evidence-Based Reading and Writing scale.[8] This transformation was possible only for students with both older PSAT scores, so only those students were retained. In a similar manner, older PSAT Mathematics scores were transformed to the revised PSAT Math scale. The same procedures were carried out for older SAT scores using the appropriate concordance table.[9]

Before merging, the PARCC summative data files were filtered for attemptedness. That is, students who were judged as not attempting at least 25% of the items on the PARCC assessment were removed from the data set. The data preparation results are summarized in Table 2.2.

---

[7] https://www.act.org/content/act/en/education-and-career-planning/college-and-career-readiness-standards/benchmarks.html
[8] https://collegereadiness.collegeboard.org/pdf/2015-psat-nmsqt-concordance-tables.pdf
[9] https://collegereadiness.collegeboard.org/pdf/higher-ed-brief-sat-concordance.pdf

Table 2.3
Sample Sizes by Test and Subject After Data Preparation

Test	Subject	Sample size
PSAT	Math	306,266
	EBRW	304,110
SAT	Math	170,577
	EBRW	170,538
ACT	English	35,918
	Reading	36,050
	Math	35,913

All of the following data merges were attempted:
- Algebra I, Geometry, and Algebra II with PSAT10 Math
- Algebra I, Geometry, and Algebra II with PSAT/NMSQT Math
- Algebra I, Geometry, and Algebra II with SAT Math
- Algebra I, Geometry, and Algebra II with ACT Math
- Grades 9, 10, and 11 ELA/L with PSAT10 Evidence-Based Reading and Writing
- Grades 9, 10, and 11 ELA/L with PSAT/NMSQT Evidence-Based Reading and Writing
- Grades 9, 10, and 11 ELA/L with SAT Evidence-Based Reading and Writing
- Grades 9, 10, and 11 ELA/L with ACT English
- Grades 9, 10, and 11 ELA/L with ACT Reading

Some combinations of assessments resulted in very small sample sizes (reported in the results section). The external assessment data were merged with PARCC summative data files (from spring 2015, fall 2015, and spring 2016) using the PARCC student identifier. Only students with a grade level when assessed (from the PARCC summative data files) between 8 and 12 were retained. Next, the period of assessment administration was identified for PARCC and the external assessments. Students were retained in the data if they took PARCC and the external assessment within approximately 6 months of each other. In addition, the sequence of the PARCC and external assessments were indicated based on whether the tests were administered during the same period, whether PARCC was administered first, or whether the external assessment was administered first. Table 2.3 lists the combinations of the PARCC and external assessments based on the sequencing and period of administration.

Table 2.4
Sequencing of Administrations for PARCC and the External Assessments

Administration Timing Indicator	PARCC Administration	External Assessment Administration
Concurrent	Spring 2015	Spring 2015
	Fall 2015	Fall 2015
	Spring 2016	Spring 2016
PARCC First	Spring 2015	Fall 2015
	Fall 2015	Spring 2016
External First	Spring 2015	Fall 2014
	Fall 2015	Spring 2015
	Spring 2016	Fall 2015

The PSAT has two sets of college readiness benchmarks: PSAT10 (generally taken in the fall of grade 10) and PSAT/NMSQT (generally taken in the fall of grade 11). For this reason, it is important to consider the grade level of students taking the PSAT. Unfortunately, grade was not consistently provided in the PSAT data. For that reason, it was necessary to infer the PSAT grade level from the PARCC grade level when assessed. When analyzing PSAT data, only students for which a given benchmark applied were retained. That is, for the PSAT10 benchmarks, only students in grade 10 or lower were analyzed. For the PSAT/NMSQT benchmarks, only students in grade 11 were analyzed.

Summarizing Demographics

To examine the representativeness of the study sample, sample demographics were compared to the PARCC population of test takers for a given assessment. The population was defined by the full summative data (from spring 2015, fall 2015, and spring 2016) set before merging with the external assessment data (after filtering for attemptedness and duplicates, including only current PARCC states). The sample and population were compared in terms of PARCC achievement and the distribution of the following demographic variables: gender, ethnicity, English language learner status, economically disadvantaged status, and students with disability status. Sample and population values were compared using an effect size (i.e., difference in standard deviation units). For proportions (percentages), the effect size was

$$\text{Cohen's } h = 2(\arcsin \sqrt{p_1} - \arcsin \sqrt{p_2}).$$

For PARCC scale scores, the effect size was Cohen's *d* with the population standard deviation as the denominator.

To better understand the sample, tables of grade level when assessed by assessment timing were generated. For example, a table like that could show that the majority of the sample is students who took a PARCC assessment in the spring of grade 9, then the external assessment

in the fall of grade 10. Alternatively, it might show that most students in the sample took the tests at the same time. Such information aids the interpretation of results.

Analysis

The College Board and ACT college readiness benchmarks are shown in Tables 2.1 and 2.2, respectively. The college readiness benchmark for all PARCC tests is 750, which is the Level 4 performance level threshold. The benchmark scores were used to create dichotomous variables indicating whether students scored below the benchmark (0) or met or exceeded the benchmark (1). The dichotomous indicator variables were used to generate 2×2 cross-tabulations showing percentages meeting neither, one, or both benchmarks. These variables would also be used later when fitting logistic regression models.

Regression

Ordinary least squares (OLS) regression estimates the linear relationship between two variables. First, external assessment scores were regressed on PARCC scale scores, then PARCC scale scores were regressed on external assessment scores. The resulting slope and intercept were used to estimate expected assessment scores. For example, if external assessment scores (Y) were regressed on PARCC scores (X), the result would be the slope (b_0) and intercept (b_1) in the linear equation

$$Y = b_0 + b_1 X.$$

Plugging the PARCC benchmark ($X = 750$) into this equation would give the expected external assessment score for a student who just met the PARCC benchmark.

The correlation between assessment scores was calculated, as was the correlation adjusted for possible restriction of range (e.g., in case the sample was less variable than the population). The adjusted correlation was calculated as

$$R = (rS/s)\sqrt{1 - r^2 + r^2 \left(\frac{S^2}{s^2}\right)},$$

where R is the unrestricted correlation, r is the restricted (sample) correlation, S is the unrestricted (population) standard deviation, and s is the restricted (sample) standard deviation.

As a non-parametric alternative to OLS regression, a locally estimated scatterplot smoothing (LOESS) procedure was applied to the data. The advantage of the LOESS fit is that it is a local estimate of the relationship between two variables, so it is not influenced by distant points. For instance, if the relationship near the benchmark is of greatest interest, that relationship is based on students near the benchmark, not those with scores far from the benchmark (e.g., students with very low scores who may not have been motivated to give their best efforts). In

this study, the LOESS span parameter (α), which controls the degree of smoothing, was set to 0.50. With α = 0.50, the neighborhood for estimation included half of the data points, with greater weight applied to closer data points.

Logistic Regression

Logistic regression models the relationship between continuous variables and the probability of being classified in one of two possible states. Here, logistic regression was used to model the relationship between students' PARCC scores, and their probabilities of meeting the external assessment benchmark (and vice versa). Fitting the logistic regression generalized linear model (GLM) estimates the intercept (b_0) and slope (b_0) in the equation

$$logit\ Y_{01} = b_0 + b_1 X \qquad \text{or}$$

$$P(Y_{01} = 1) = 1/(1 + e^{-(b_0 + b_1 X)}),$$

where Y_{01} is the dichotomous (0/1) variable indicating whether a student met or exceeded the benchmark score. Plugging the PARCC benchmark ($X = 750$) into this equation would give the probability of meeting or exceeding the external assessment benchmark for a student who just met the PARCC benchmark. The non-parametric LOESS fit was also used to estimate probabilities of meeting or exceeding benchmarks (using span α = 0.50).

3. Results

In the plots provided in this section, data points are color coded by state; however the specific state is masked (State A = black, State B = red, State C = blue, State D = green, and State E = orange) to illustrate the sources of the data and to diagnose possible causes of unexpected results. Several comparisons had insufficient samples to complete; these are reviewed in Appendix B.

3.1 Algebra I
3.1.1 Algebra I and PSAT10 Math

The full sample had lower average ability than the population of PARCC Algebra I test takers (Table 3.1). Other notable differences included a higher percentage of Black or African American students and a lower percentage of Hispanic students. As indicated in Table 3.2, the majority of students in this sample (52.5%) took Algebra I as ninth graders, then took the PSAT in the fall of tenth grade. The next largest group (27.6%) took the PSAT in the fall of ninth grade, then took Algebra I in the spring of ninth grade.

Table 3.1

Demographic Comparison for Sample Taking PARCC Algebra I and PSAT10 Math

Variable	PARCC Population	Sample	Difference	Effect Size
N	624,008	42,637		
% Female	51.7	50.1	-1.6	-0.03
% American Indian/Alaskan Native	1.2	0.7	-0.5	-0.05
% Asian	6.0	4.0	-2.0	-0.09
% Black/African American	18.3	39.7	21.5	0.48
% Hispanic	28.3	16.4	-11.9	-0.29
% Native Hawaiian/Other Pacific Islander	0.2	0.2	0.0	0.00
% White	42.2	35.6	-6.6	-0.14
% English Language Learners	7.3	6.2	-1.0	-0.04
% Economically Disadvantaged	43.2	47.7	4.4	0.09
% Students with Disability	12.5	12.1	-0.4	-0.01
% PARCC Level 4 or Level 5	31.3	18.3	-13.0	-0.30
Scale Score Mean	733.6	725.0	-8.6	-0.26
Scale Score SD	33.7	28.2		

Table 3.2

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.0	0.2	0.5	0.0	0.0
External First	0.5	27.6	10.9	0.0	0.0
PARCC First	7.8	52.5	0.0	0.0	0.0

As shown in Table 3.3, the percentages of students meeting the PSAT10 benchmark (16.0%) was similar to the percentage of students meeting Level 4 or higher on Algebra I (18.3%). Students who met Level 4 or higher on Algebra I were expected to score slightly below the PSAT10 benchmark of 480 (Table 3.4 and Figure 3.1). Likewise, students who met the PSAT10 benchmark were expected to score slightly below the PARCC benchmark of 750. The probabilities of meeting the benchmarks based on logistic regression support the same conclusions (Table 3.4 and Figure 3.2). Students who scored 750 on PARCC Algebra I had a 0.285 probability of meeting the PSAT10 Math benchmark. The adjusted correlation between scores was .612.

Table 3.3

Cross-Tabulation Percentages for Sample Taking PARCC Algebra I and PSAT10 Math

	Below Level 4	Level 4 or Level 5	
Below Benchmark	75.2	8.9	84.0
Met Benchmark	6.5	9.4	16.0
	81.7	18.3	

Table 3.4

Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Algebra I and PSAT10 Math

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Algebra I	750	42,637	437.8	435.3	.285	.265
PSAT10 Math	480	42,637	740.9	742.1	.403	.394
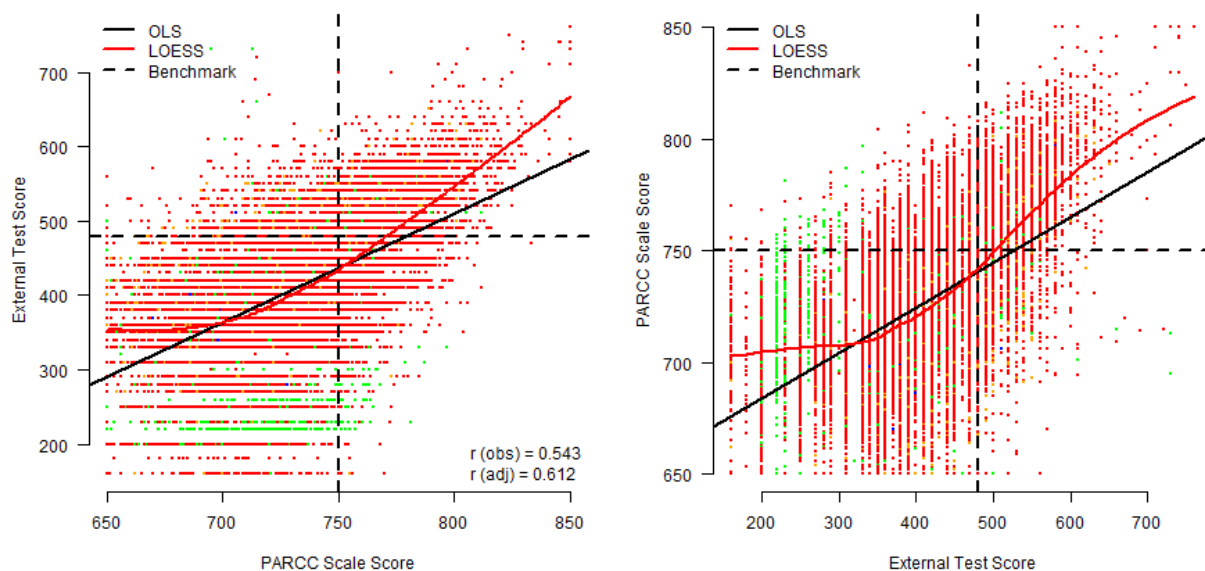
Figure 3.1. OLS and LOESS regression for Algebra I and PSAT10 Math.
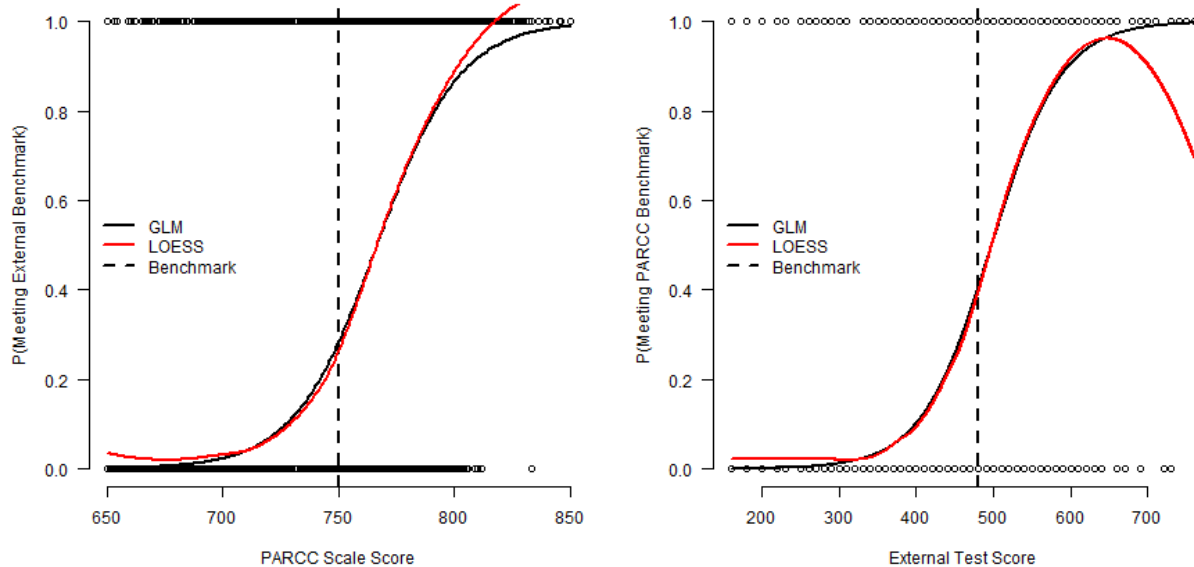
Figure 3.2. Logistic regression for Algebra I and PSAT10 Math.

To investigate possible differences between students who took PARCC first and who took the PSAT first, the analysis was re-run using only students who took PARCC Algebra I in ninth grade and PSAT in tenth grade. The overall results were generally similar (Table 3.5), with students meeting the benchmark for one assessment expected to score below the benchmark on the other assessment.

Table 3.5
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Algebra I and PSAT10 Math (Only Students Taking PARCC in Ninth Grade and PSAT10 in Tenth Grade)

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Algebra I	750	22,397	442.5	441.3	.328	.300
PSAT10 Math	480	22,397	734.5	735.9	.284	.281

3.2 Geometry
3.2.1 Geometry and PSAT10 Math
The Geometry/PSAT10 Math sample was similar in average ability to the population of PARCC Geometry test takers (Table 3.6). The largest difference between sample and population was that the sample included a greater percentage of Hispanic students. The overwhelming majority of students in this sample (87.1%) took the PSAT in the fall of tenth grade, then took PARCC Geometry in the spring of tenth grade (Table 3.7).

Table 3.6

Demographic Comparison for Sample Taking PARCC Geometry and PSAT10 Math

Variable	PARCC Population	Sample	Difference	Effect Size
N	289,864	15,779		
% Female	51.3	49.1	-2.2	-0.04
% American Indian/Alaskan Native	1.9	6.8	4.9	0.25
% Asian	7.0	3.3	-3.7	-0.17
% Black/African American	12.7	5.5	-7.2	-0.25
% Hispanic	29.3	43.0	13.7	0.29
% Native Hawaiian/Other Pacific Islander	0.2	0.1	-0.1	-0.01
% White	46.4	40.1	-6.4	-0.13
% English Language Learners	5.2	6.8	1.7	0.07
% Economically Disadvantaged	38.4	47.1	8.7	0.18
% Students with Disability	13.0	9.5	-3.4	-0.11
% PARCC Level 4 or Level 5	23.9	18.8	-5.2	-0.13
Scale Score Mean	730.0	727.7	-2.3	-0.09
Scale Score SD	26.7	24.5		

Table 3.7

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.1	0.0	2.8	0.0	0.0
External First	0.0	1.3	87.1	0.0	0.0
PARCC First	0.0	8.6	0.0	0.0	0.0

As shown in Table 3.8, the percentage of students meeting the PSAT10 benchmark (33.7%) was higher than the percentage of students meeting Level 4 or higher on Geometry (18.8%). Students who met Level 4 or higher on Geometry were expected to exceed the PSAT10 benchmark of 480 (Table 3.8 and Figure 3.3). Students who met the PSAT10 benchmark were expected to score below the PARCC benchmark of 750. The probabilities of meeting the benchmarks based on logistic regression support the same conclusions (Table 3.9 and Figure 3.4). Students who scored 750 on PARCC Geometry had a 0.665 probability of having met the PSAT10 Math benchmark near the start of tenth grade. Had these students taken the tests concurrently (in the spring of grade 10), their average PSAT10 performance would likely be improved, so the percentages of students meeting the benchmarks would likely have been closer. The adjusted correlation between scores was 0.663.

Table 3.8

Cross-Tabulation Percentages for Sample Taking PARCC Geometry and PSAT10 Math

	Below Level 4	Level 4 or Level 5	
Below Benchmark	63.4	2.9	66.3
Met Benchmark	17.8	15.8	33.7
	81.2	18.8	

Table 3.9

Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Geometry and PSAT10 Math

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Geometry	750	15,779	486.8	486.8	.665	.668
PSAT10 Math	480	15,779	735.5	733.4	.203	.169
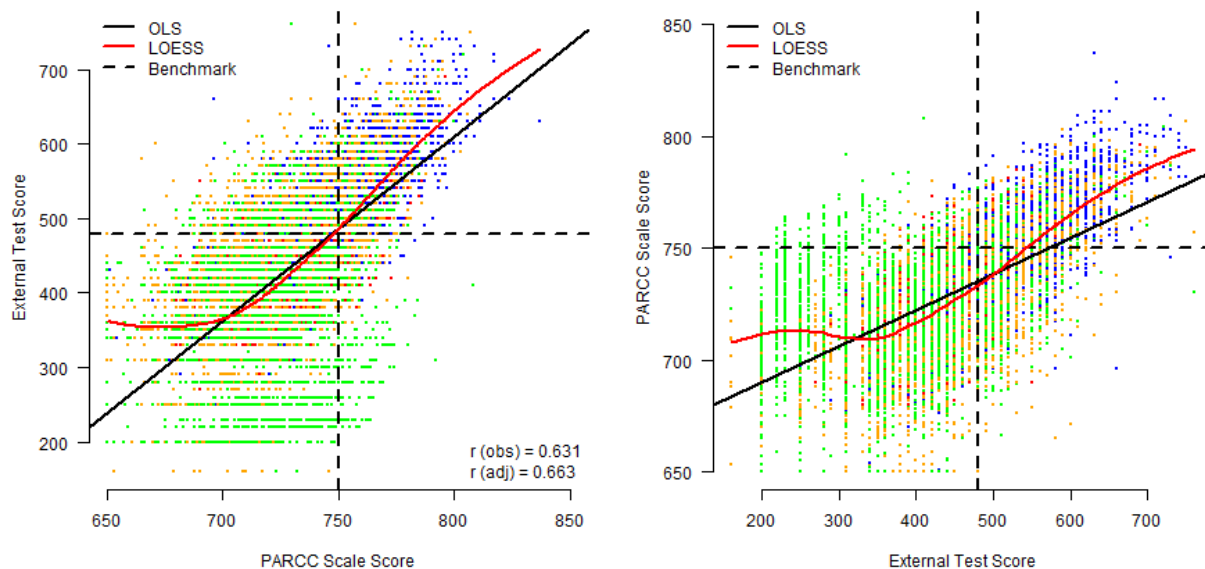
Figure 3.3. OLS and LOESS regression for Geometry and PSAT10 Math.
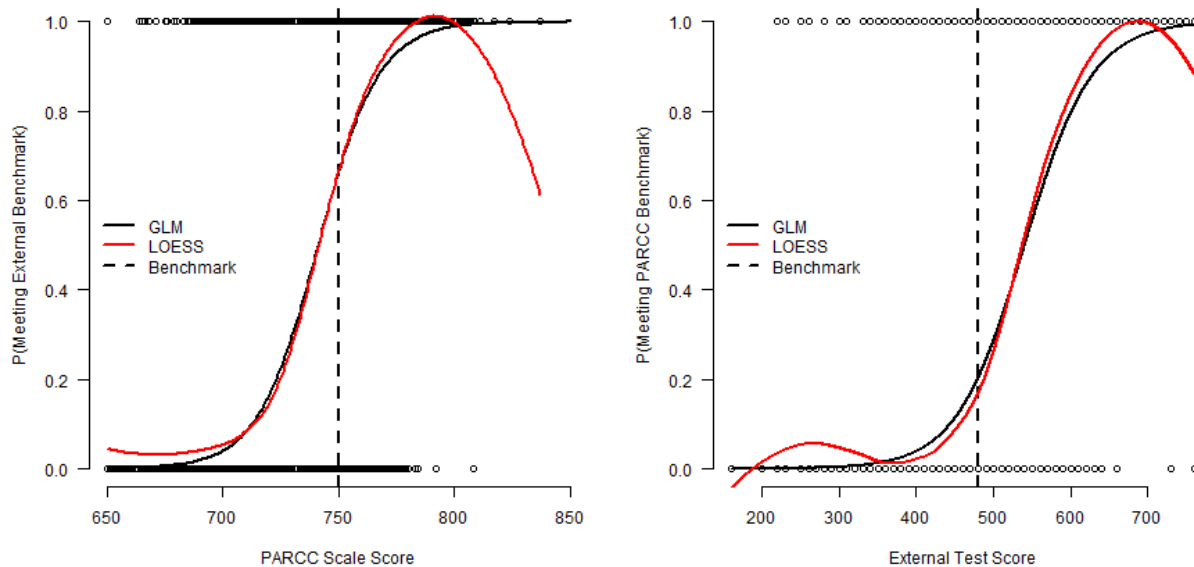
Figure 3.4. Logistic regression for Algebra I and PSAT10 Math.

3.3 Algebra II
3.3.1 Algebra II and PSAT10 Math

The Algebra II/PSAT10 Math sample was notably higher in average ability than the population of PARCC Algebra II test takers (Table 3.10). Compared to the population, the sample included a smaller percentage of Hispanic students and economically disadvantaged students. The majority of students in this sample (71.4%) took the PSAT in the fall of tenth grade, then took PARCC Algebra II in the spring of tenth grade (Table 3.11).

Table 3.10
Demographic Comparison for Sample Taking PARCC Algebra II and PSAT10 Math

Variable	PARCC Population	Sample	Difference	Effect Size
N	322,956	51,259		
% Female	49.7	47.8	-2.0	-0.04
% American Indian/Alaskan Native	1.5	0.6	-0.9	-0.09
% Asian	7.4	14.3	6.9	0.23
% Black/African American	15.5	17.2	1.8	0.05
% Hispanic	23.6	12.2	-11.4	-0.30
% Native Hawaiian/Other Pacific Islander	0.2	0.2	0.0	0.00
% White	48.6	53.1	4.4	0.09
% English Language Learners	3.1	0.9	-2.1	-0.16
% Economically Disadvantaged	35.2	20.9	-14.3	-0.32

Variable	PARCC Population	Sample	Difference	Effect Size
% Students with Disability	8.3	2.5	-5.9	-0.27
% PARCC Level 4 or Level 5	22.4	40.0	17.6	0.38
Scale Score Mean	720.1	739.3	19.2	0.51
Scale Score SD	37.9	35.6		

Table 3.11

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.0	0.1	1.4	0.0	0.0
External First	0.2	6.7	71.4	0.0	0.0
PARCC First	0.8	19.6	0.0	0.0	0.0

As shown in Table 3.12, the percentages of students meeting the PSAT10 benchmark (74.8%) was higher than the percentage of students meeting Level 4 or higher on Algebra II (40.0%). Students who met Level 4 or higher on Algebra II were expected to exceed the PSAT10 benchmark of 480 (Table 3.12 and Figure 3.5). Students who met the PSAT10 benchmark were expected to score below the PARCC benchmark of 750. The probabilities of meeting the benchmarks based on logistic regression support the same conclusions (Table 3.13 and Figure 3.6). Students who scored 750 on PARCC Algebra II had a 0.886 probability of having met the PSAT10 Math benchmark near the start of tenth grade. Had these students taken the tests concurrently (in the spring of grade 10), their average PSAT10 performance would likely be improved, so the percentages of students meeting the benchmarks would likely have been closer. The adjusted correlation between scores was 0.731.

Table 3.12

Cross-Tabulation Percentages for Sample Taking PARCC Algebra II and PSAT10 Math

	Below Level 4	Level 4 or Level 5	
Below Benchmark	23.9	1.4	25.2
Met Benchmark	36.1	38.6	74.8
	60.0	40.0	

Table 3.13

Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Algebra II and PSAT10 Math

| Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Algebra II	750	51,259	549.2	541.0	.886	.894
PSAT10 Math	480	51,259	726.1	723.4	.151	.135

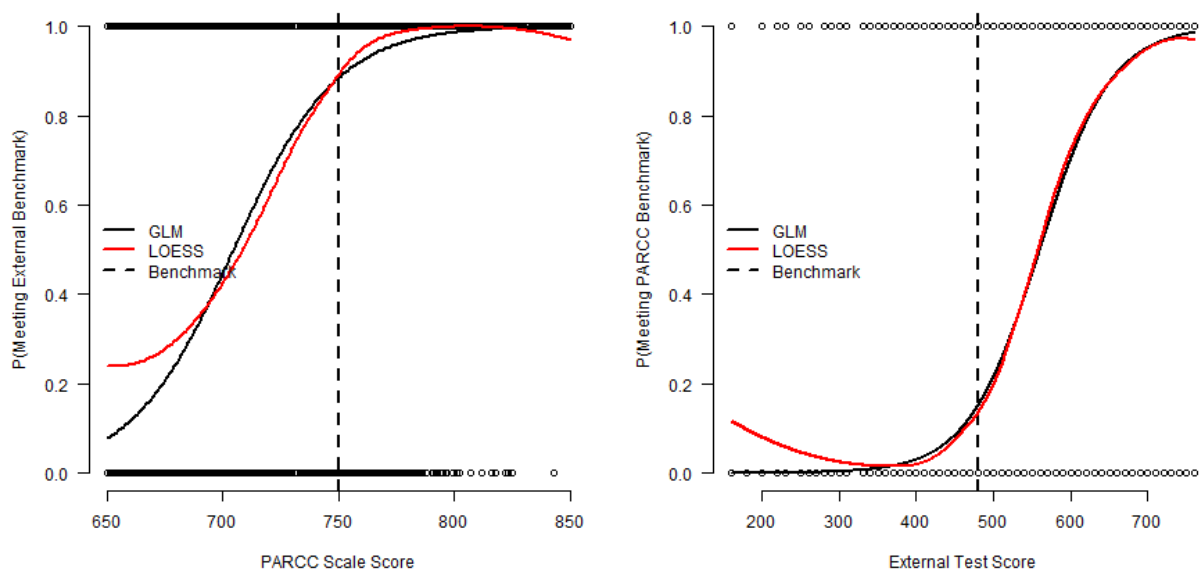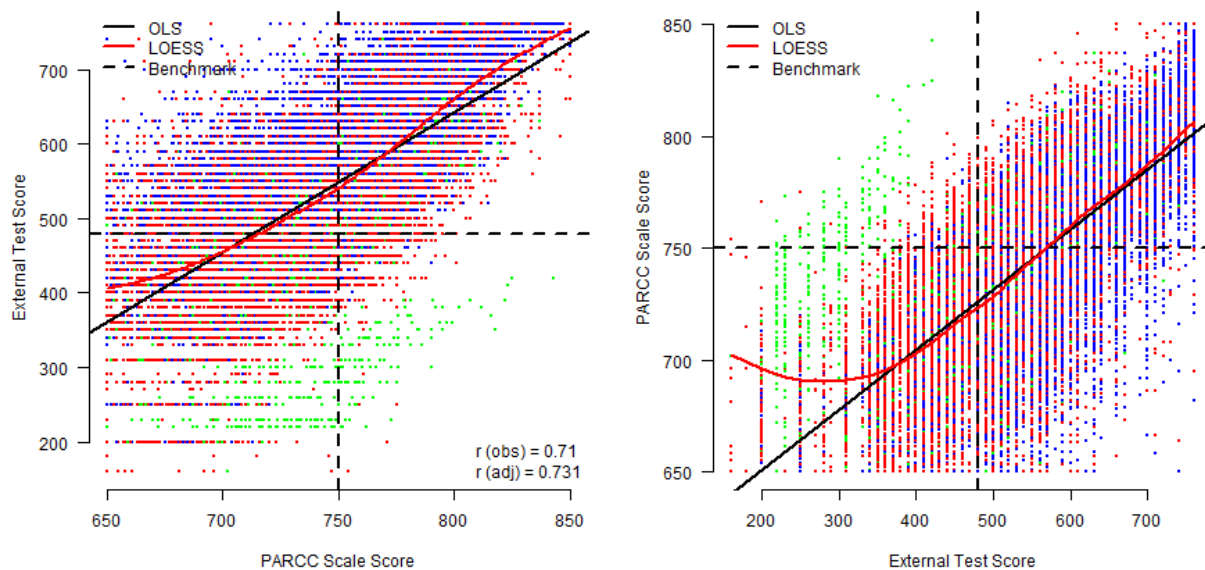*Figure 3.5*. OLS and LOESS regression for Algebra II and PSAT10 Math.

Figure 3.6. Logistic regression for Algebra II and PSAT10 Math.

In Figure 3.5, there appears to be a cloud of aberrant points that do not follow the general relationship between Algebra II and PSAT10 scores (green points with unusually low PSAT10 performance). To investigate the influence of these points on results, the analysis was re-run

without data from State D. As would be expected, the expected PSAT10 score for students who scored 750 on PARCC Algebra II increased, as did the probability of meeting the PSAT10 benchmark (Table 3.14). However, the overall effect was small since there were only 3,287 students from State D in the data.

Table 3.14
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Algebra II and PSAT10 Math (Excluding One State's Data)

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Algebra II	750	47,972	552.0	544.1	.895	.901
PSAT10 Math	480	47,972	724.8	723.1	.140	.133

3.3.2 Algebra II and PSAT/NMSQT Math

There was a large sample of students who took PARCC Algebra II and the PSAT/NMSQT within six months of each other. The sample had similar average ability to the population of PARCC Algebra II testers, and the sample was fairly representative of the PARCC population in terms of demographics (Table 3.15). The sample was a mix of students who took the PSAT/NMSQT in the fall of eleventh grade and PARCC Algebra II the following spring and students who took PARCC Algebra II as tenth graders and the PSAT/NMSQT in the fall as eleventh graders (Table 3.16).

Table 3.15
Demographic Comparison for Sample Taking PARCC Algebra II and PSAT/NMSQT Math

Variable	PARCC Population	Sample	Difference	Effect Size
N	322,956	62,775		
% Female	49.7	47.9	-1.9	-0.04
% American Indian/Alaskan Native	1.5	0.4	-1.1	-0.13
% Asian	7.4	9.0	1.6	0.06
% Black/African American	15.5	21.5	6.0	0.16
% Hispanic	23.6	17.8	-5.8	-0.14
% Native Hawaiian/Other Pacific Islander	0.2	0.2	0.0	0.00
% White	48.6	49.5	0.8	0.02
% English Language Learners	3.1	1.7	-1.4	-0.09
% Economically Disadvantaged	35.2	30.1	-5.1	-0.11
% Students with Disability	8.3	7.7	-0.7	-0.02
% PARCC Level 4 or Level 5	22.4	19.6	-2.9	-0.07
Scale Score Mean	720.1	718.7	-1.4	-0.04
Scale Score SD	37.9	35.2		

Table 3.16

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.0	0.0	0.0	0.9	0.1
External First	0.0	0.0	0.0	54.6	0.4
PARCC First	0.0	0.0	41.7	2.4	0.0

As shown in Table 3.17, a higher percentage of students met the PSAT/NMSQT benchmark than met the PARCC benchmark (44.1% vs. 19.6%). Students who met Level 4 or higher on Algebra II were expected to exceed the PSAT/NMSQT benchmark of 510 (Table 3.17 and Figure 3.7). Students who met the PSAT10 benchmark were expected to score below the PARCC benchmark of 750. The probabilities of meeting the benchmarks based on logistic regression support the same conclusions (Table 3.18 and Figure 3.8). Students who scored 750 on PARCC Algebra II had a 0.760 probability of meeting the PSAT/NMSQT Math benchmark. The adjusted correlation between scores was 0.718.

Table 3.17

Cross-Tabulation Percentages for Sample Taking PARCC Algebra II and
PSAT/NMSQT Math

	Below Level 4	Level 4 or Level 5	
Below Benchmark	54.4	1.5	55.9
Met Benchmark	26.1	18.1	44.1
	80.4	19.6	

Table 3.18

Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Algebra II and
PSAT/NMSQT Math

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Algebra II	750	62,775	553.8	551.5	.760	.790
PSAT Math	510	62,775	722.8	718.9	.113	.105

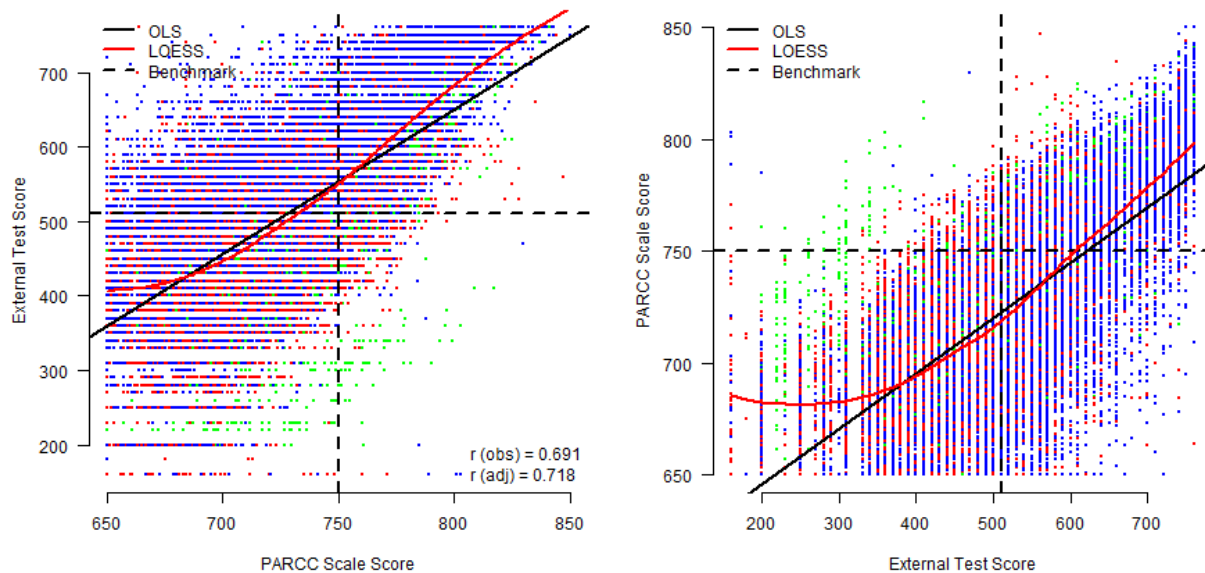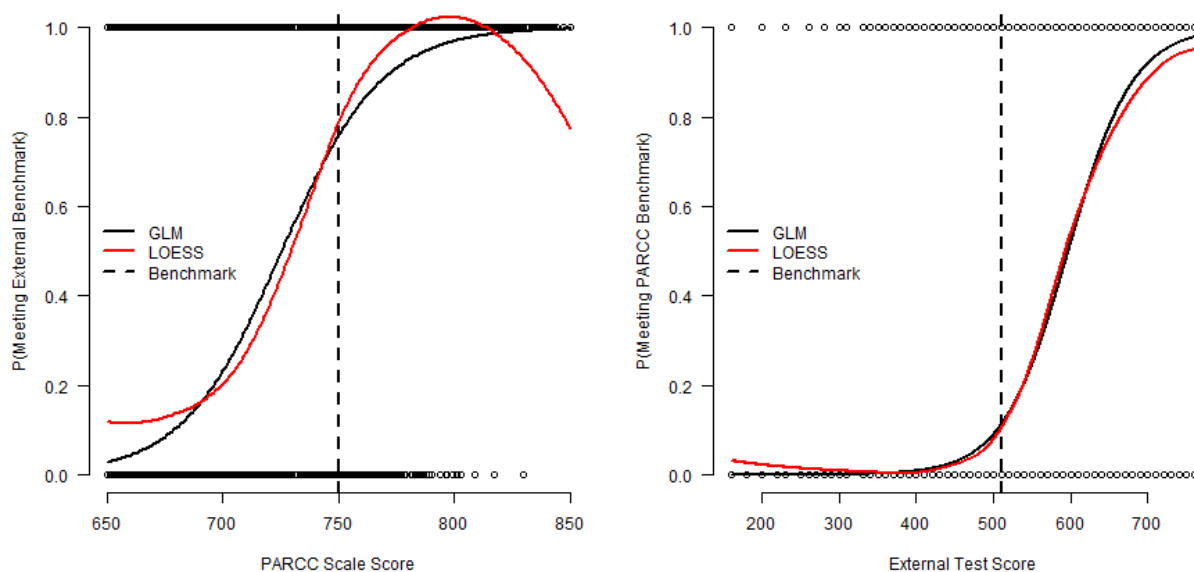*Figure 3.7*. OLS and LOESS regression for Algebra II and PSAT/NMSQT Math.

Figure 3.8. Logistic regression for Algebra II and PSAT/NMSQT Math.

Because the sample of test takers was split so evenly between students who took PARCC Algebra II first and those who took the PSAT/NMSQT first, the analyses were re-run separately for those two groups. In both cases, the general trend in results was the same: students

meeting the PARCC benchmark were expected to exceed the PSAT/NMSQT benchmark, and students meeting the PSAT/NMSQT benchmark were expected to score well below the PARCC benchmark. The major difference in results was that students who took Algebra II as tenth graders before taking the PSAT/NMSQT were expected to score significantly higher on the PSAT/NMSQT (Table 3.19) than students who took the tests in the reverse order (Table 3.20). This difference is likely related to sample differences in average ability, since students taking Algebra II as tenth graders are of higher average ability than students taking Algebra II as eleventh graders.

Table 3.19
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Algebra II and PSAT/NMSQT Math (Only Students Who Took Algebra II Before PSAT/NMSQT)

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Algebra II	750	27,638	572.3	565.7	.840	.855
PSAT Math	510	27,638	725.4	722.4	.127	.112

Table 3.20
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Algebra II and PSAT/NMSQT Math (Only Students Who Took PSAT/NMSQT Before Algebra II)

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Algebra II	750	34,509	522.2	527.3	.636	.677
PSAT Math	510	34,509	717.9	716.5	.098	.098

3.3.3 Algebra II and SAT Math

There was a large sample of students who took PARCC Algebra II and the SAT within six months of each other. The sample was of lower average ability and included more Black students than the PARCC population of Algebra II test takers (Table 3.21). The sample included a mix of students who took both tests around the same time as eleventh graders or who took PARCC Algebra II as an eleventh grader then the SAT the following fall.

Table 3.21
Demographic Comparison for Sample Taking PARCC Algebra II and SAT Math

Variable	PARCC Population	Sample	Difference	Effect Size
N	322,956	59,154		
% Female	49.7	45.2	-4.5	-0.09
% American Indian/Alaskan Native	1.5	0.2	-1.3	-0.16
% Asian	7.4	7.7	0.3	0.01
% Black/African American	15.5	25.1	9.6	0.24
% Hispanic	23.6	19.5	-4.1	-0.10
% Native Hawaiian/Other Pacific Islander	0.2	0.2	0.0	0.01

Variable	PARCC Population	Sample	Difference	Effect Size
% White	48.6	41.5	-7.1	-0.14
% English Language Learners	3.1	2.7	-0.3	-0.02
% Economically Disadvantaged	35.2	34.0	-1.2	-0.03
% Students with Disability	8.3	7.8	-0.5	-0.02
% PARCC Level 4 or Level 5	22.4	13.0	-9.4	-0.25
Scale Score Mean	720.1	711.7	-8.4	-0.22
Scale Score SD	37.9	33.0		

Table 3.22

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.0	0.2	1.9	47.5	0.9
External First	0.0	0.0	0.1	1.4	2.5
PARCC First	0.0	0.2	7.7	37.4	0.0

As shown in Table 3.22, a higher percentage of students met the SAT Math benchmark than met the PARCC benchmark (33.2% vs. 13.0%). Students who met Level 4 or higher on Algebra II were expected to exceed the SAT benchmark of 530 (Table 3.23 and Figure 3.9). Students who met the SAT benchmark were expected to score below the PARCC benchmark of 750. The probabilities of meeting the benchmarks based on logistic regression support the same conclusions (Table 3.24 and Figure 3.10). Students who scored 750 on PARCC Algebra II had a 0.697 probability of meeting the SAT Math benchmark. The adjusted correlation between scores was 0.711.

Table 3.23

Cross-Tabulation Percentages for Sample Taking PARCC Algebra II and SAT Math

	Below Level 4	Level 4 or Level 5	
Below Benchmark	65.5	1.3	66.8
Met Benchmark	21.5	11.7	33.2
	87.0	13.0	

Table 3.24

Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Algebra II and SAT Math

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Algebra II	750	59,154	563.8	565.2	.697	.740
SAT Math	530	59,154	720.8	718.7	.097	.107

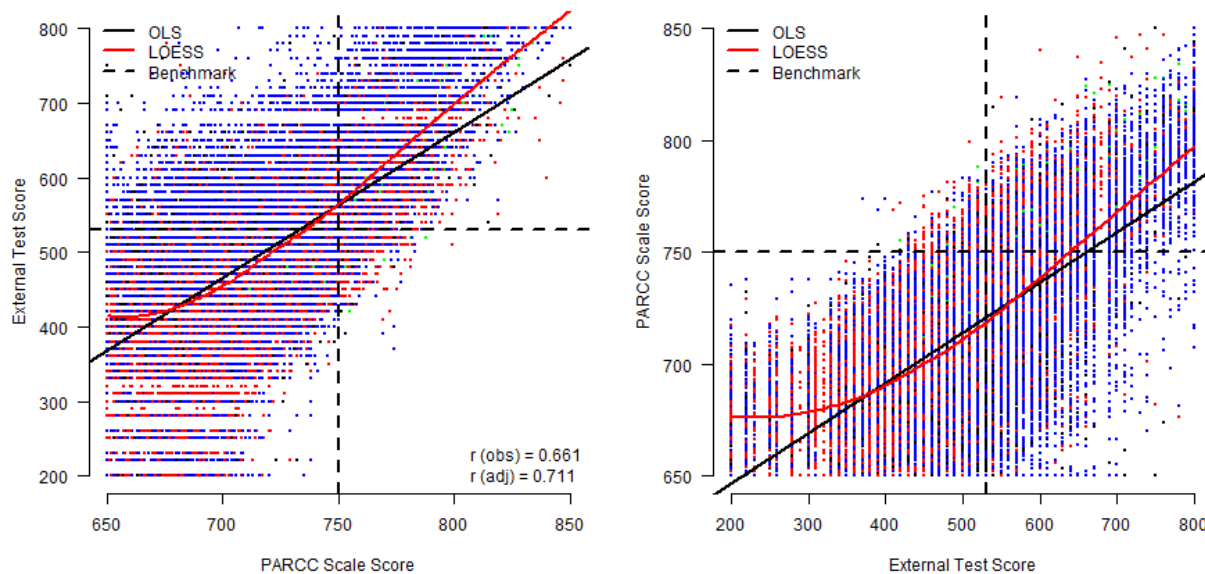*Figure 3.9*. OLS and LOESS regression for Algebra II and SAT Math.
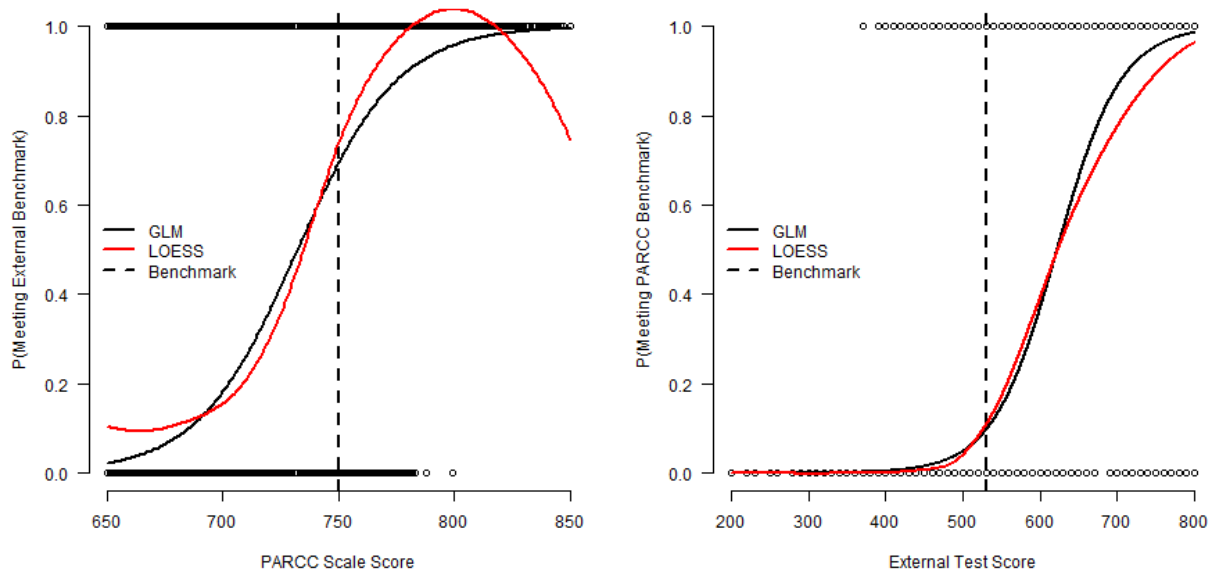

Figure 3.10. Logistic regression for Algebra II and SAT Math.

Because the sample of test takers was split between students who took PARCC Algebra II and SAT concurrently and those who took PARCC Algebra II before the SAT, the analyses were re-run separately for those two groups. In both cases, the general trend in results was the same:

students meeting the PARCC benchmark were expected to exceed the SAT benchmark, and students meeting the SAT benchmark were expected to score well below the PARCC benchmark. The major difference in results was that students who took Algebra II before the SAT (Table 3.25) were expected to score significantly higher on the SAT than students who took the tests concurrently (in eleventh grade; Table 3.26). This result might be expected since students might be expected to perform better on the SAT in the fall of grade 12 than in the spring of grade 11.

Table 3.25
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Algebra II and SAT Math (Only Students Who Took Algebra II Concurrently With SAT)

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Algebra II	750	29,914	550.6	555.4	.648	.689
SAT Math	510	29,914	720.9	719.7	.103	.116

Table 3.26
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Algebra II and SAT Math (Only Students Who Took Algebra II Before SAT)

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Algebra II	750	26,851	576.9	575.9	.747	.794
SAT Math	510	26,851	720.7	717.8	.090	.096

3.3.4 Algebra II and ACT Math

The available sample of students who took PARCC Algebra II and the ACT within 6 months of each other was representative of the PARCC Algebra II population in terms of ability, but the sample included fewer females, more Hispanic students, and fewer White students than the population (Table 3.27). The sample was split between students who took the two test concurrently as eleventh graders and students who took PARCC Algebra II in grade 11 then the ACT in grade 12 (Table 3.28).

Table 3.27

Demographic Comparison for Sample Taking PARCC Algebra II and ACT Math

Variable	PARCC Population	Sample	Difference	Effect Size
N	322,956	6,627		
% Female	49.7	39.5	-10.2	-0.21
% American Indian/Alaskan Native	1.5	6.8	5.3	0.28
% Asian	7.4	2.9	-4.5	-0.21
% Black/African American	15.5	15.4	-0.1	0.00
% Hispanic	23.6	37.8	14.2	0.31
% Native Hawaiian/Other Pacific Islander	0.2	0.1	-0.1	-0.02
% White	48.6	35.4	-13.3	-0.27
% English Language Learners	3.1	2.6	-0.4	-0.03
% Economically Disadvantaged	35.2	40.3	5.1	0.11
% Students with Disability	8.3	4.0	-4.3	-0.18
% PARCC Level 4 or Level 5	22.4	17.8	-4.6	-0.12
Scale Score Mean	720.1	721.1	1.0	0.03
Scale Score SD	37.9	31.6		

Table 3.28

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.0	0.2	1.8	40.3	5.3
External First	0.0	0.1	0.3	6.6	6.1
PARCC First	0.0	0.1	3.7	33.5	2.1

As shown in Table 3.28, a higher percentage of students met the ACT Math benchmark than met the PARCC benchmark (23.8% vs. 17.8%). Students who earned a score of 750 on PARCC Algebra II were expected to score very close to the ACT Math benchmark of 22 (Table 3.29 and Figure 3.11). In contrast, students who just met the ACT Math benchmark were expected to score below the PARCC benchmark of 750. That pattern of results is corroborated by the logistic regression results (Table 3.30 and Figure 3.12). Students who scored 750 on PARCC Algebra II had 0.509 probability of meeting the ACT benchmark. The adjusted correlation between scores was 0.767.

Table 3.29

Cross-Tabulation Percentages for Sample Taking PARCC Algebra II and ACT Math

	Below Level 4	Level 4 or Level 5	
Below Benchmark	72.4	3.8	76.2
Met Benchmark	9.7	14.1	23.8
	82.2	17.8	

Table 3.30

Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Algebra II and ACT Math

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Algebra II	750	6,627	21.3	21.2	.509	.491
ACT Math	22	6,627	739.4	737.6	.310	.320
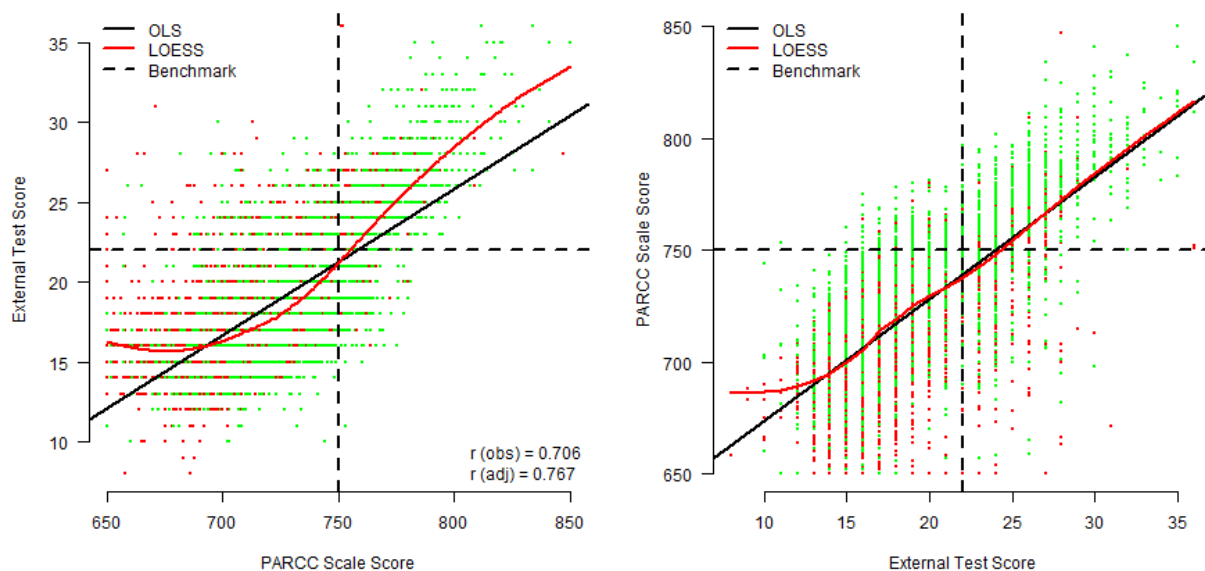
Figure 3.11. OLS and LOESS regression for Algebra II and ACT Math.
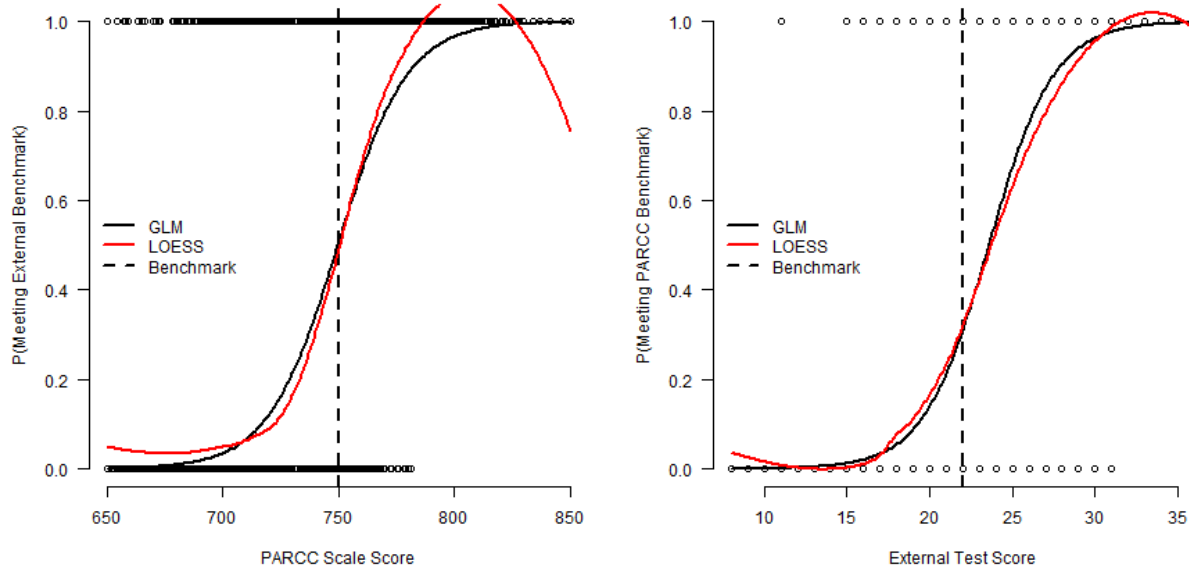
Figure 3.12. Logistic regression for Algebra II and ACT Math.

Because the sample of test takers was split between students who took PARCC Algebra II and ACT concurrently and those who took PARCC Algebra II before the ACT, the analyses were re-run separately for those two groups. In both cases, the general trend in results was the same: students meeting the PARCC benchmark were expected to score close to the ACT benchmark, and students meeting the ACT benchmark were expected to score below the PARCC benchmark. The major difference in results was that students who took the ACT earlier (concurrent with Algebra II in grade 11; Table 3.31) had higher expected performance on Algebra II than students who took the ACT in the fall of grade 12 (Table 3.32). This result might be expected since students might be expected to perform better on the ACT in the fall of grade 12 than in the spring of grade 11. This result might be expected since the "concurrent" group was of slightly higher average ability (725 vs. 717 average Algebra II scale score).

Table 3.31
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Algebra II and ACT Math (Only Students Who Took Algebra II Concurrently With ACT)

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Algebra II	750	3,150	21.4	21.0	.504	.471
ACT Math	22	3,150	742.1	739.3	.374	.367

Table 3.32
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Algebra II and ACT Math (Only Students Who Took Algebra II Before ACT)

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Algebra II	750	2,610	21.3	21.9	.543	.565
ACT Math	22	2,610	734.7	734.0	.222	.231

3.4 Grade 9 ELA/L
3.4.1 Grade 9 ELA/L and PSAT10 EBRW

The available sample of students who took PARCC Grade 9 ELA/L and the PSAT10 Evidence-Based Reading and Writing test was significantly higher in average ability than the PARCC Grade 9 ELA/L population. Compared to the population, the sample included more Asian students and fewer economically disadvantaged students (Table 3.33). Most of the sample (91.5%) was students who took Grade 9 ELA/L as ninth graders and PSAT10 in the fall of grade 10 (Table 3.34).

Table 3.33

Demographic Comparison for Sample Taking PARCC Grade 9 ELA/L and PSAT10 EBRW

Variable	PARCC Population	Sample	Difference	Effect Size
N	531,082	8,836		
% Female	51.4	49.0	-2.4	-0.05
% American Indian/Alaskan Native	1.4	2.1	0.7	0.06
% Asian	6.1	18.6	12.6	0.40
% Black/African American	13.6	5.1	-8.6	-0.30
% Hispanic	31.0	21.1	-9.9	-0.23
% Native Hawaiian/Other Pacific Islander	0.2	0.3	0.1	0.03
% White	44.2	51.4	7.1	0.14
% English Language Learners	6.7	3.5	-3.3	-0.15
% Economically Disadvantaged	42.7	28.7	-14.0	-0.29
% Students with Disability	11.8	4.3	-7.5	-0.28
% PARCC Level 4 or Level 5	38.0	65.3	27.2	0.55
Scale Score Mean	737.6	762.5	24.9	0.67
Scale Score SD	37.1	39.6		

Table 3.34

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.0	0.1	0.4	0.0	0.0
External First	0.0	7.4	0.5	0.0	0.0
PARCC First	0.0	91.5	0.0	0.0	0.0

As shown in Table 3.35, a slightly higher percentage of students met the PSAT10 EBRW benchmark than met the PARCC benchmark (71.1% vs. 65.3%). Students who earned a score of 750 on PARCC Grade 9 ELA/L were expected to exceed the PSAT10 benchmark score of 430 (Table 3.35 and Figure 3.13). Students who just met the PSAT10 benchmark were expected to score very close to the PARCC benchmark score of 750 (747.9). That pattern of results is

consistent with the logistic regression analysis, which indicated that students who scored 750 on PARCC Grade 9 ELA/L had a 0.684 probability of meeting the PSAT10 benchmark, and students who just met the PSAT10 benchmark had a 0.513 probability of meeting the PARCC benchmark (Table 3.36 and Figure 3.14). The adjusted correlation between scores was 0.652.

Table 3.35
Cross-Tabulation Percentages for Sample Taking PARCC Grade 9 ELA/L and PSAT10 EBRW

	Below Level 4	Level 4 or Level 5	
Below Benchmark	21.0	7.9	28.9
Met Benchmark	13.7	57.3	71.1
	34.7	65.3	

Table 3.36
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Grade 9 ELA/L and PSAT10 Evidence-Based Reading and Writing

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Grade 9 ELA/L	750	8,836	471.1	449.6	.684	.662
PSAT10 EBRW	430	8,836	747.9	737.4	.513	.353
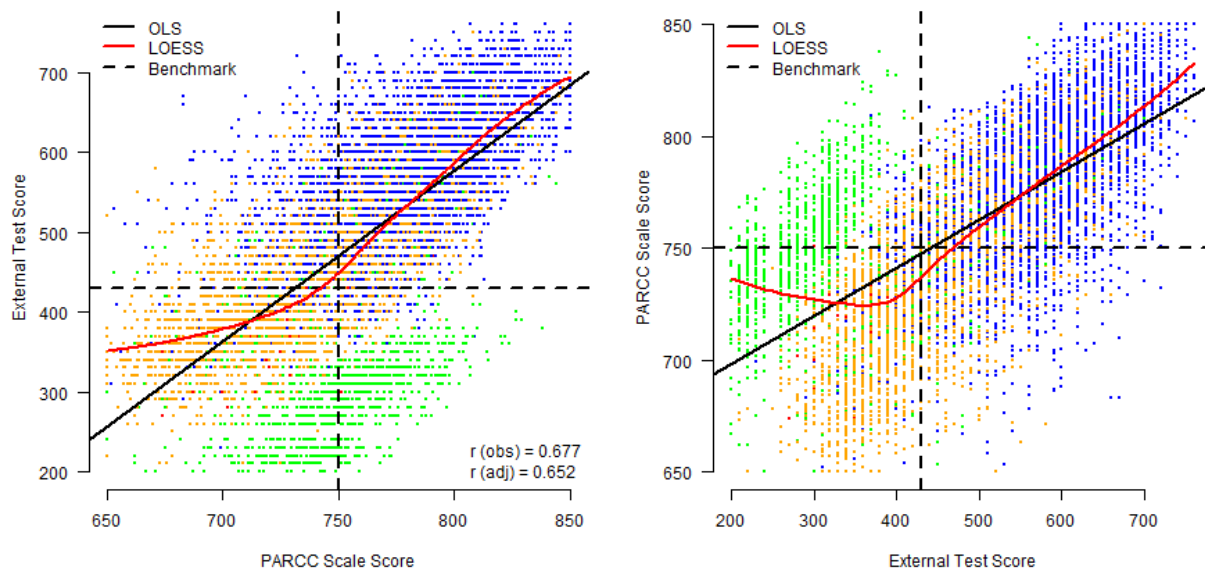
Figure 3.13. OLS and LOESS regression for Grade 9 ELA/L and PSAT10 EBRW.
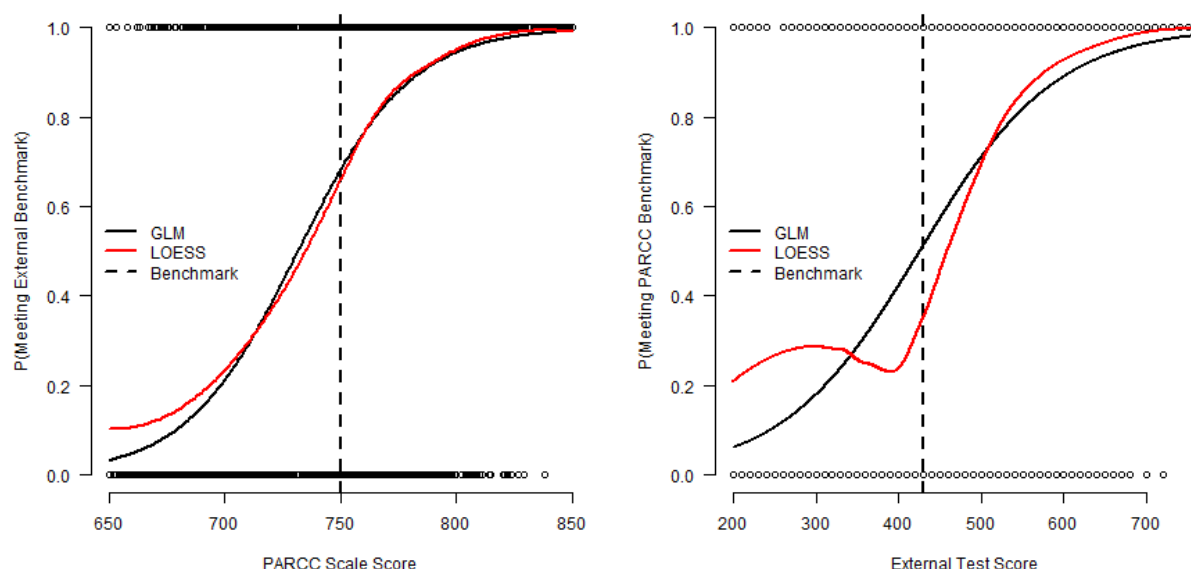
Figure 3.14. Logistic regression for Grade 9 ELA/L and PSAT10 EBRW.

In Figure 3.13, there appears to be a cloud of aberrant points that do not follow the general relationship between Grade 9 ELA/L and PSAT10 scores (green points with unusually low PSAT10 performance). To investigate the influence of these points on results, the analysis was re-run without data State D. As would be expected, the expected PSAT10 score for students who scored 750 on PARCC Grade 9 ELA/L increased, as did the probability of meeting the PSAT10 benchmark (Table 3.37).

Table 3.37
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Grade 9 ELA/L and PSAT10 EBRW (Excluding One State's Data)

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Grade 9 ELA/L	750	7,438	496.9	494.8	.842	.831
PSAT10 EBRW	430	7,438	735.6	737.5	.330	.361

3.5 Grade 10 ELA/L
3.5.1 Grade 10 ELA/L and PSAT10
There was a large sample of students who took PARCC Grade 10ELA/L and the PSAT10 Evidence-Based Reading and Writing (Table 3.38). This sample had higher average ability than the PARCC Grade 10 ELA/L population, and it included fewer Hispanic students. Nearly all of

the sample (96.5%) was students who took PSAT10 in the fall of grade 10 and PARCC Grade 10 ELA/L the following spring (Table 3.39).

Table 3.38

Demographic Comparison for Sample Taking PARCC Grade 10 ELA/L and PSAT10 EBRW

Variable	PARCC Population	Sample	Difference	Effect Size
N	395,803	122,251		
% Female	51.3	48.9	-2.4	-0.05
% American Indian/Alaskan Native	1.6	1.3	-0.3	-0.02
% Asian	6.8	9.0	2.2	0.08
% Black/African American	19.3	25.6	6.3	0.15
% Hispanic	25.8	15.9	-9.9	-0.25
% Native Hawaiian/Other Pacific Islander	0.2	0.1	0.0	-0.01
% White	44.1	45.2	1.1	0.02
% English Language Learners	4.9	2.0	-2.8	-0.16
% Economically Disadvantaged	39.1	31.9	-7.2	-0.15
% Students with Disability	13.2	8.1	-5.1	-0.17
% PARCC Level 4 or Level 5	38.6	50.3	11.7	0.24
Scale Score Mean	735.8	749.3	13.5	0.29
Scale Score SD	46.0	45.5		

Table 3.39

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.0	0.0	1.9	0.0	0.0
External First	0.0	1.0	96.5	0.0	0.0
PARCC First	0.0	0.6	0.0	0.0	0.0

As shown in Table 3.40, a higher percentage of students met the PSAT10 EBRW benchmark than met the PARCC benchmark (61.5% vs. 50.3%). Students who earned a score of 750 on PARCC Grade 10 ELA/L were expected to exceed the PSAT10 benchmark score of 430 (Table 3.40 and Figure 3.15). Students who just met the PSAT10 benchmark were expected to score 737.7, which is below the PARCC benchmark score of 750. That pattern of results is consistent with the logistic regression analysis, which indicated that students who scored 750 on PARCC Grade 10 ELA/L had a 0.692 probability of meeting the PSAT10 benchmark, and students who just met the PSAT10 benchmark had a 0.357 probability of meeting the PARCC benchmark (Table 3.41 and Figure 3.16). The adjusted correlation between scores was 0.733.

Table 3.40

Cross-Tabulation Percentages for Sample Taking PARCC Grade 10 ELA/L and PSAT10 EBRW

	Below Level 4	Level 4 or Level 5	
Below Benchmark	33.1	5.4	38.5
Met Benchmark	16.6	44.9	61.5
	49.7	50.3	

Table 3.41

Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Grade 10 ELA/L and PSAT10 Evidence-Based Reading and Writing

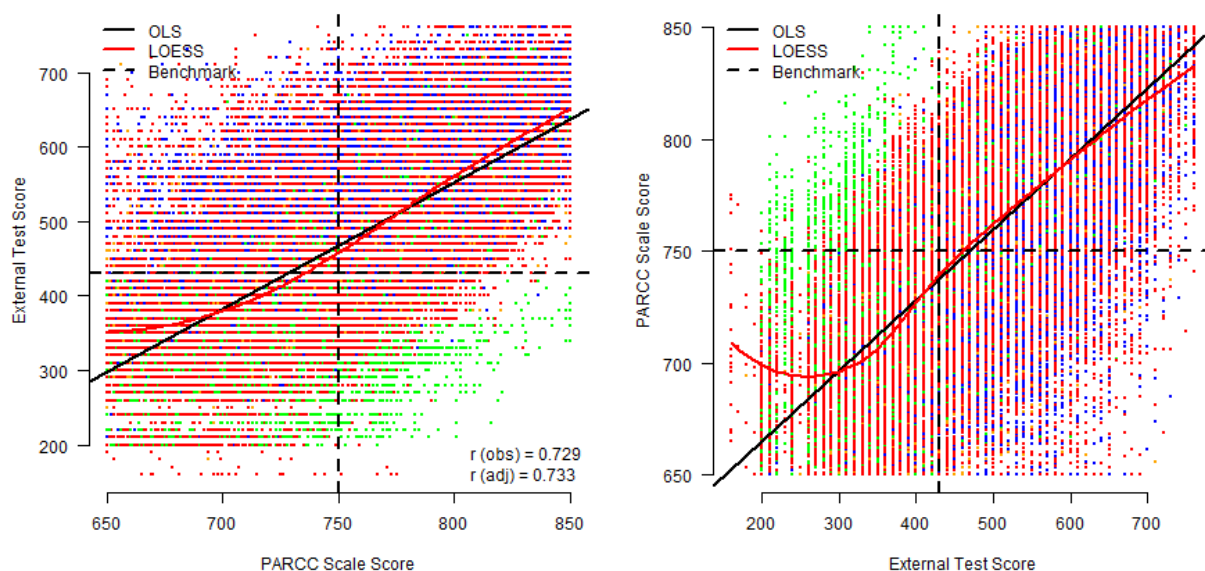| Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Grade 10						
ELA/L	750	122,251	468.0	458.3	.692	.667
PSAT10 EBRW	430	122,251	737.7	739.5	.357	.380

Figure 3.15. OLS and LOESS regression for Grade 10 ELA/L and PSAT10 EBRW.
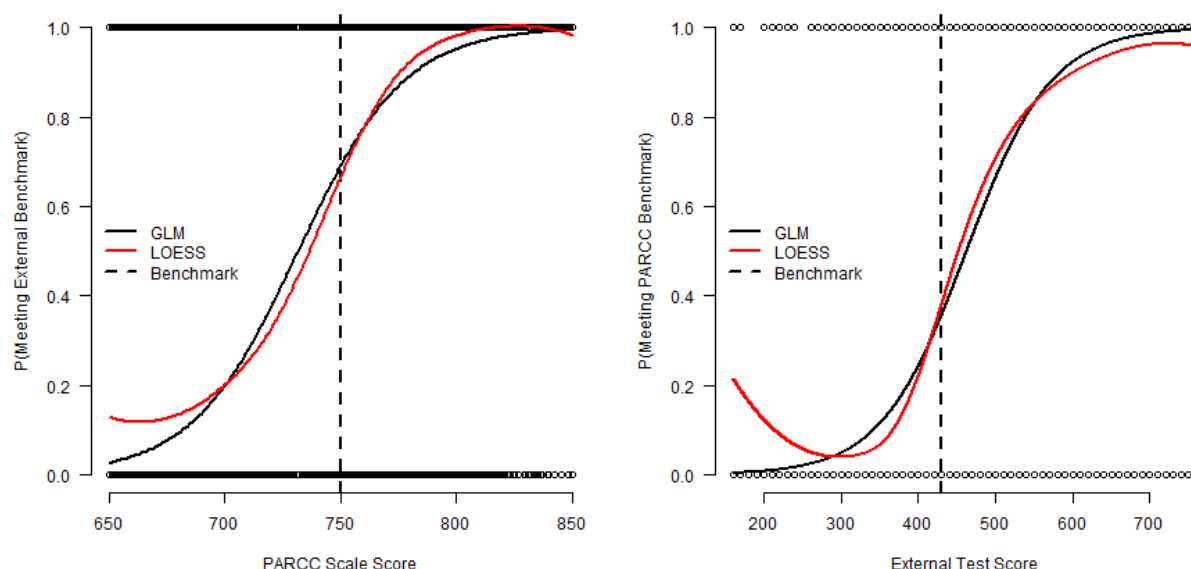
Figure 3.16. Logistic regression for Grade 10 ELA/L and PSAT10 EBRW.

Again with the PSAT10 data, there was a cloud of aberrant points that did not follow the general relationship between Grade 10 ELA/L and PSAT10 scores (green points in Figure 3.15 with unusually low PSAT10 performance). To investigate the influence of these points on results, the analysis was re-run without data from State D. As would be expected, the expected PSAT10 score for students who scored 750 on PARCC Grade 10 ELA/L increased, as did the probability of meeting the PSAT10 benchmark (Table 3.42). The effect was small since State D only contributed 12,050 to the overall sample size.

Table 3.42
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Grade 10 ELA/L and PSAT10 EBRW (Excluding One State's Data)

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Grade 10 ELA/L	750	110,201	472.2	463.0	.718	.687
PSAT10 EBRW	430	110,201	736.5	739.2	.344	.380

3.5.2 Grade 10 ELA/L and PSAT/NMSQT Evidence Based Reading and Writing

The available data included 47,170 students who took PARCC Grade 10 ELA/L and the PSAT/NMSQT Evidence-Based Reading and Writing test (Table 3.43). This sample had higher average ability than the PARCC Grade 10 ELA/L population, and it included fewer Hispanic students and fewer economically disadvantaged students. Nearly all of the sample (97.3%) was

students who took PARCC Grade 10 ELA/L as tenth graders and the PSAT/NMSQT in the fall of eleventh grade (Table 3.44).

Table 3.43

Demographic Comparison for Sample Taking PARCC Grade 10 ELA/L and PSAT/NMSQT EBRW

Variable	PARCC Population	Sample	Difference	Effect Size
N	395,803	47,170		
% Female	51.3	48.6	-2.7	-0.05
% American Indian/Alaskan Native	1.6	0.2	-1.4	-0.17
% Asian	6.8	11.9	5.1	0.18
% Black/African American	19.3	23.0	3.6	0.09
% Hispanic	25.8	10.9	-14.9	-0.39
% Native Hawaiian/Other Pacific Islander	0.2	0.1	0.0	-0.01
% White	44.1	51.1	7.0	0.14
% English Language Learners	4.9	1.5	-3.3	-0.20
% Economically Disadvantaged	39.1	26.0	-13.1	-0.28
% Students with Disability	13.2	5.2	-8.0	-0.28
% PARCC Level 4 or Level 5	38.6	53.2	14.7	0.30
Scale Score Mean	735.8	752.5	16.7	0.36
Scale Score SD	46.0	44.2		

Table 3.44

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.0	0.0	0.0	0.1	0.0
External First	0.0	0.0	0.0	2.5	0.1
PARCC First	0.0	0.0	97.3	0.1	0.0

As shown in Table 3.45, a higher percentage of students met the PSAT/NMSQT EBRW benchmark than met the PARCC benchmark (70.4% vs. 53.2%). Students who earned a score of 750 on PARCC Grade 10 ELA/L were expected to exceed the PSAT/NMSQT benchmark score of 460 by more than 50 points (Table 3.45 and Figure 3.17). Students who just met the PSAT/NMSQT benchmark were expected to score 736.0, which is below the PARCC benchmark score of 750. Logistic regression results showed the same general pattern, with students who scored 750 on PARCC Grade 10 ELA/L having a 0.775 probability of meeting the PSAT/NMSQT benchmark, and students who just met the PSAT/NMSQT benchmark having a 0.326 probability of meeting the PARCC benchmark (Table 3.46 and Figure 3.18). The adjusted correlation between scores was 0.723.

Table 3.45

Cross-Tabulation Percentages for Sample Taking PARCC Grade 10 ELA/L and PSAT/NMSQT EBRW

	Below Level 4	Level 4 or Level 5	
Below Benchmark	25.7	3.9	29.6
Met Benchmark	21.1	49.3	70.4
	46.8	53.2	

Table 3.46

Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Grade 10 ELA/L and PSAT/NMSQT Evidence-Based Reading and Writing

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Grade 10 ELA/L	750	47,170	511.2	504.6	.775	.756
PSAT EBRW	460	47,170	736.0	736.9	.326	.346

Figure 3.17. OLS and LOESS regression for Grade 10 ELA/L and PSAT/NMSQT EBRW.
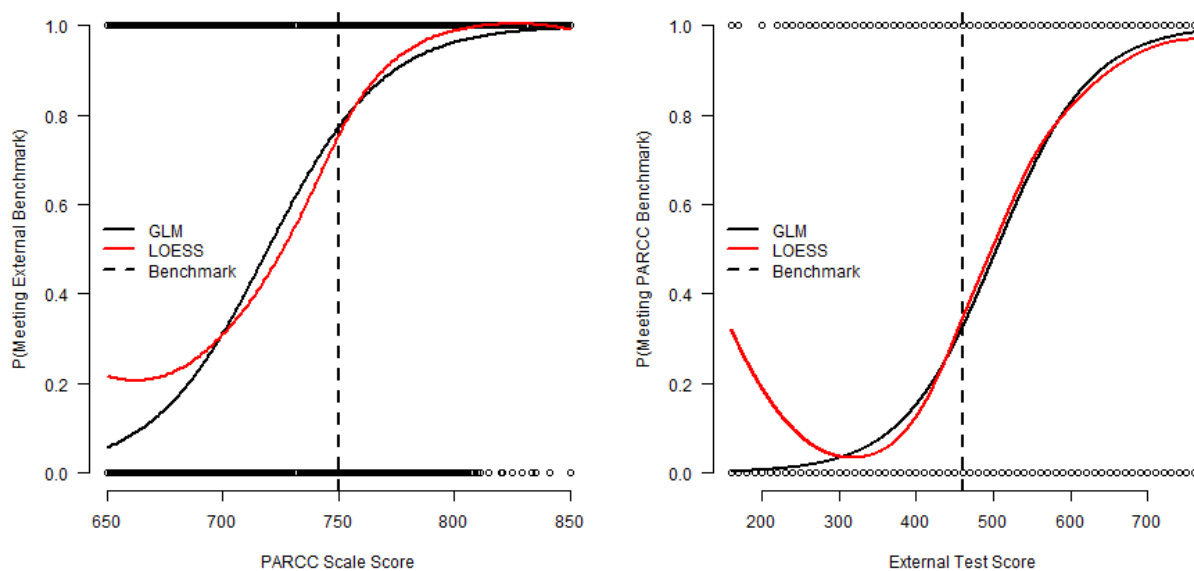
Figure 3.18. Logistic regression for Grade 10 ELA/L and PSAT/NMSQT EBRW.

3.5.3 Grade 10 ELA/L and SAT Evidence Based Reading and Writing

There was a very high ability sample of students who took PARCC Grade 10 ELA/L and the PSAT/NMSQT Evidence-Based Reading and Writing test (Table 3.47). This sample had more Asian and White students and fewer Hispanic students, economically disadvantaged students, and students with disabilities than the population of PARCC Grade 10 ELA/L test takers. The majority of these students (63.0%) took PARCC Grade 10 ELA/L as tenth graders and the SAT in the fall of eleventh grade (Table 3.48).

Table 3.47

Demographic Comparison for Sample Taking PARCC Grade 10 ELA/L and SAT EBRW

Variable	PARCC Population	Sample	Difference	Effect Size
N	395,803	10,195		
% Female	51.3	46.2	-5.1	-0.10
% American Indian/Alaskan Native	1.6	0.1	-1.5	-0.18
% Asian	6.8	20.2	13.4	0.40
% Black/African American	19.3	15.5	-3.8	-0.10
% Hispanic	25.8	7.3	-18.5	-0.52
% Native Hawaiian/Other Pacific Islander	0.2	0.1	-0.1	-0.02
% White	44.1	54.3	10.2	0.21
% English Language Learners	4.9	4.7	-0.1	-0.01
% Economically Disadvantaged	39.1	16.9	-22.2	-0.50

Variable	PARCC Population	Sample	Difference	Effect Size
% Students with Disability	13.2	2.5	-10.8	-0.43
% PARCC Level 4 or Level 5	38.6	65.1	26.5	0.54
Scale Score Mean	735.8	765.0	29.1	0.63
Scale Score SD	46.0	47.4		

Table 3.48

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.0	0.1	21.4	7.0	0.8
External First	0.0	0.0	2.4	0.5	2.6
PARCC First	0.0	0.1	63.0	1.9	0.2

A higher percentage of students met the SAT EBRW benchmark than met the PARCC benchmark (79.5% vs. 65.1%; Table 3.49). Students who earned a score of 750 on PARCC Grade 10 ELA/L were expected to exceed the SAT benchmark score of 480 by more than 60 points (Table 3.49 and Figure 3.19). Students who just met the SAT benchmark were expected to score 738.5, which is below the PARCC benchmark score of 750. Logistic regression indicated that students who scored 750 on PARCC Grade 10 ELA/L had a 0.825 probability of meeting the SAT benchmark, and students who just met the SAT benchmark had a 0.385 probability of meeting the PARCC benchmark (Table 3.50 and Figure 3.20). The adjusted correlation between scores was 0.704.

Table 3.49
Cross-Tabulation Percentages for Sample Taking PARCC Grade 10 ELA/L and SAT EBRW

	Below Level 4	Level 4 or Level 5	
Below Benchmark	17.6	3.0	20.5
Met Benchmark	17.3	62.1	79.5
	34.9	65.1	

Table 3.50
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Grade 10 ELA/L and SAT Evidence-Based Reading and Writing

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Grade 10 ELA/L	750	10,195	546.2	545.6	.825	.785
SAT EBRW	480	10,195	738.5	741.6	.385	.426

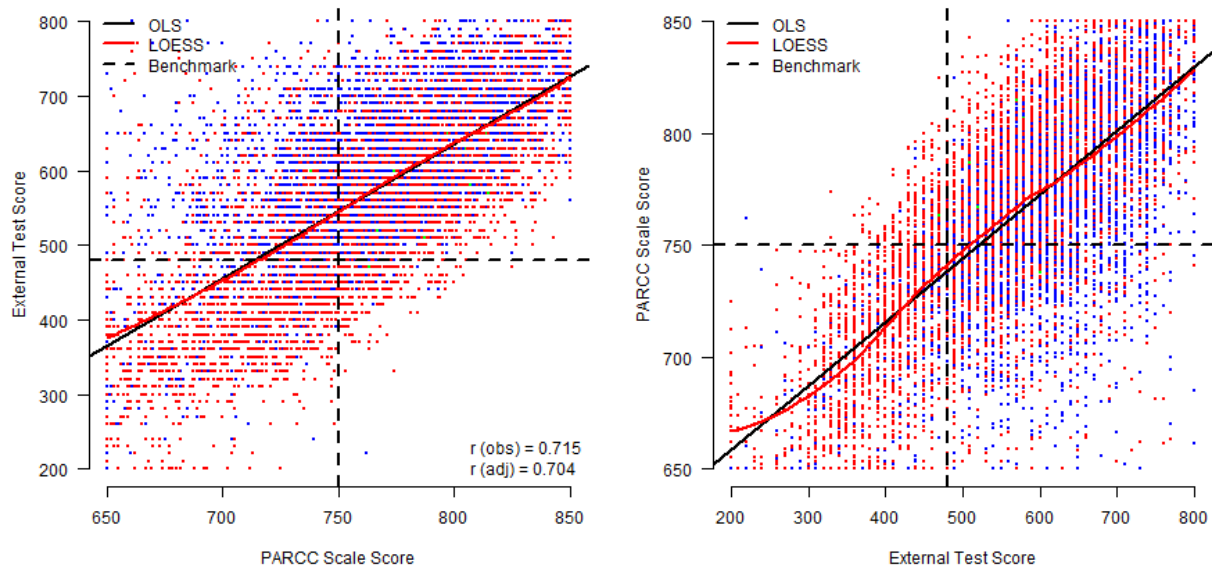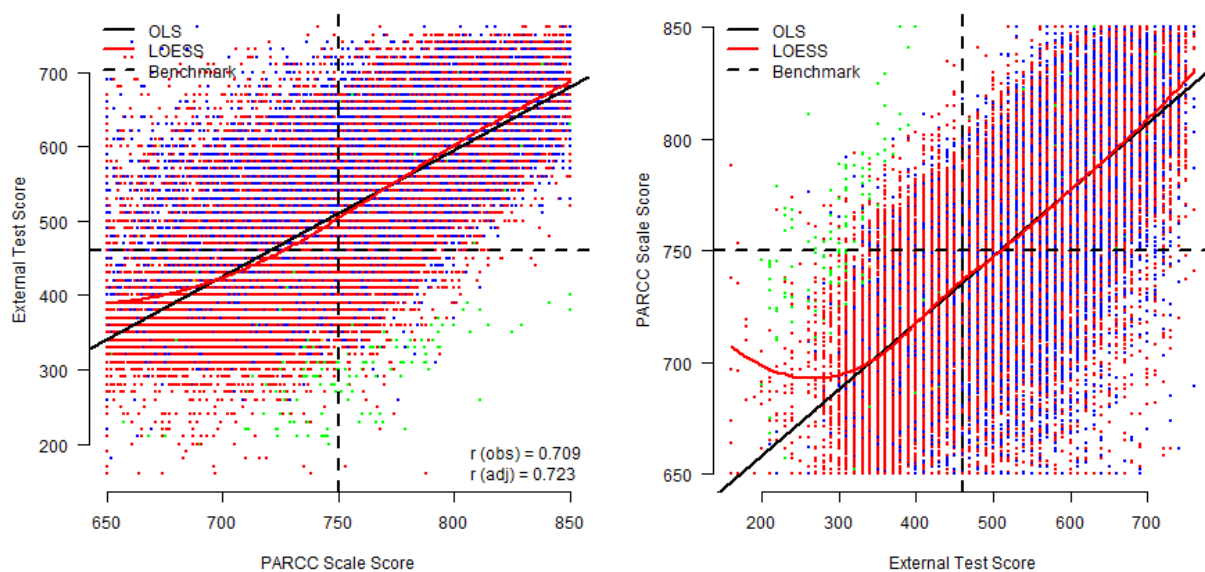*Figure 3.19*. OLS and LOESS regression for Grade 10 ELA/L and SAT EBRW.
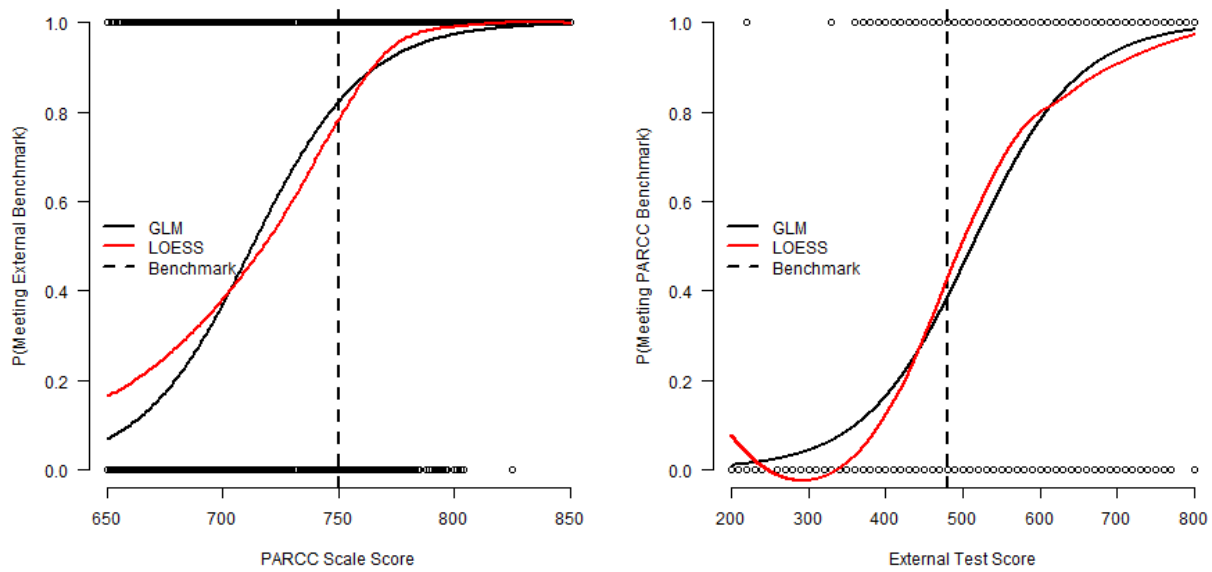

Figure 3.20. Logistic regression for Grade 10 ELA/L and SAT EBRW.

3.6 Grade 11 ELA/L
3.6.2 Grade 11 ELA/L and PSAT/NMSQT Evidence Based Reading and Writing

There was a relatively high ability sample of students who took PARCC Grade 11 ELA/L and the PSAT/NMSQT Evidence-Based Reading and Writing test (Table 3.51). This sample had more Asian students and fewer Hispanic and economically disadvantaged students than the PARCC population of Grade 11 ELA/L test takers. Most students in the sample (93.0%) took the PSAT/NMSQT in the fall of eleventh grade and PARCC Grade 11 ELA/L the following spring (Table 3.52).

Table 3.51

Demographic Comparison for Sample Taking PARCC Grade 11 ELA/L and PSAT/NMSQT EBRW

Variable	PARCC Population	Sample	Difference	Effect Size
N	304,214	64,823		
% Female	51.6	49.9	-1.8	-0.04
% American Indian/Alaskan Native	1.9	0.4	-1.6	-0.16
% Asian	5.2	10.8	5.6	0.21
% Black/African American	14.3	16.1	1.8	0.05
% Hispanic	27.0	19.5	-7.5	-0.18
% Native Hawaiian/Other Pacific Islander	0.2	0.2	0.1	0.01
% White	47.7	51.8	4.2	0.08
% English Language Learners	3.9	1.8	-2.1	-0.13
% Economically Disadvantaged	40.0	28.4	-11.5	-0.24
% Students with Disability	12.3	10.0	-2.2	-0.07
% PARCC Level 4 or Level 5	38.8	50.9	12.1	0.24
Scale Score Mean	738.2	750.2	12.0	0.31
Scale Score SD	38.7	39.7		

Table 3.52

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.0	0.0	0.0	3.1	0.0
External First	0.0	0.0	0.0	93.0	0.1
PARCC First	0.0	0.0	0.2	3.7	0.0

A higher percentage of students met the PSAT/NMSQT EBRW benchmark than met the PARCC benchmark (62.3% vs. 50.9%; Table 3.53). Students who earned a score of 750 on PARCC Grade 11 ELA/L were expected to exceed the PSAT/NMSQT benchmark score of 460 by approximately 40 points (Table 3.53 and Figure 3.21). Students who just met the SAT benchmark were expected to score 741.2, which is below the PARCC benchmark score of 750. Logistic regression

indicated that students who scored 750 on PARCC Grade 11 ELA/L had a 0.667 probability of meeting the SAT benchmark, and students who just met the SAT benchmark had a 0.395 probability of meeting the PARCC benchmark (Table 3.54 and Figure 3.22). Had these students taken the tests concurrently (in the spring of grade 11), their average PSAT/NMSQT performance would likely be improved, so the percentages of students meeting the benchmarks would likely have been closer. The adjusted correlation between scores was 0.635.

Table 3.53

Cross-Tabulation Percentages for Sample Taking PARCC Grade 11 ELA/L and PSAT/NMSQT EBRW

	Below Level 4	Level 4 or Level 5	
Below Benchmark	29.9	7.8	37.7
Met Benchmark	19.2	43.1	62.3
	49.1	50.9	

Table 3.54

Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Grade 11 ELA/L and PSAT/NMSQT Evidence-Based Reading and Writing

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Grade 11 ELA/L	750	64,823	497.6	484.6	.667	.607
PSAT/NMSQT EBRW	460	64,823	741.2	742.0	.395	.431

Figure 3.21. OLS and LOESS regression for Grade 11 ELA/L and PSAT/NMSQT EBRW.

Figure 3.22. Logistic regression for Grade 11 ELA/L and PSAT/NMSQT EBRW.

3.6.3 Grade 11 ELA/L and SAT Evidence Based Reading and Writing

There was a relatively high ability sample of students who took PARCC Grade 11 ELA/L and the SAT Evidence-Based Reading and Writing test (Table 3.55). This sample had more Asian students and fewer Hispanic students, economically disadvantaged students, and students with disabilities than the PARCC population of Grade 11 ELA/L test takers. The sample was split between students who took the tests concurrently as eleventh graders (53.4%) and students who took the PARCC assessment as eleventh graders and the SAT in the fall of twelfth grade (40.1%; Table 3.56).

Table 3.55

Demographic Comparison for Sample Taking PARCC Grade 11 ELA/L and SAT EBRW

Variable	PARCC Population	Sample	Difference	Effect Size
N	304,214	82,201		
% Female	51.6	47.2	-4.4	-0.09
% American Indian/Alaskan Native	1.9	0.1	-1.8	-0.21
% Asian	5.2	11.2	6.0	0.22
% Black/African American	14.3	15.3	1.0	0.03
% Hispanic	27.0	18.8	-8.2	-0.20
% Native Hawaiian/Other Pacific Islander	0.2	0.2	0.1	0.01
% White	47.7	47.6	-0.1	0.00
% English Language Learners	3.9	2.1	-1.8	-0.11
% Economically Disadvantaged	40.0	27.8	-12.2	-0.26
% Students with Disability	12.3	6.8	-5.5	-0.19
% PARCC Level 4 or Level 5	38.8	52.7	13.9	0.28
Scale Score Mean	738.2	751.7	13.5	0.35
Scale Score SD	38.7	38.6		

Table 3.56

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.0	0.0	0.2	53.4	0.1
External First	0.0	0.0	0.0	5.8	0.2
PARCC First	0.0	0.0	0.2	40.1	0.0

A higher percentage of students met the SAT EBRW benchmark than met the PARCC benchmark (70.5% vs. 52.7%; Table 3.57). Students who earned a score of 750 on PARCC Grade 10 ELA/L were expected to exceed the SAT benchmark score of 480 by approximately 60 points (Table 3.57 and Figure 3.23). Students who just met the SAT benchmark were expected to score 739.1, which is below the PARCC benchmark score of 750. Logistic regression indicated that students who scored 750 on PARCC Grade 10 ELA/L had a 0.750 probability of meeting the SAT benchmark, and students who just met the SAT benchmark had a 0.365 probability of meeting the PARCC benchmark (Table 3.58 and Figure 3.24). The adjusted correlation between scores was 0.629.

Table 3.57

Cross-Tabulation Percentages for Sample Taking PARCC Grade 11 ELA/L and SAT EBRW

	Below Level 4	Level 4 or Level 5	
Below Benchmark	23.8	5.7	29.5
Met Benchmark	23.5	47.1	70.5
	47.3	52.7	

Table 3.58

Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Grade 11 ELA/L and SAT Evidence-Based Reading and Writing

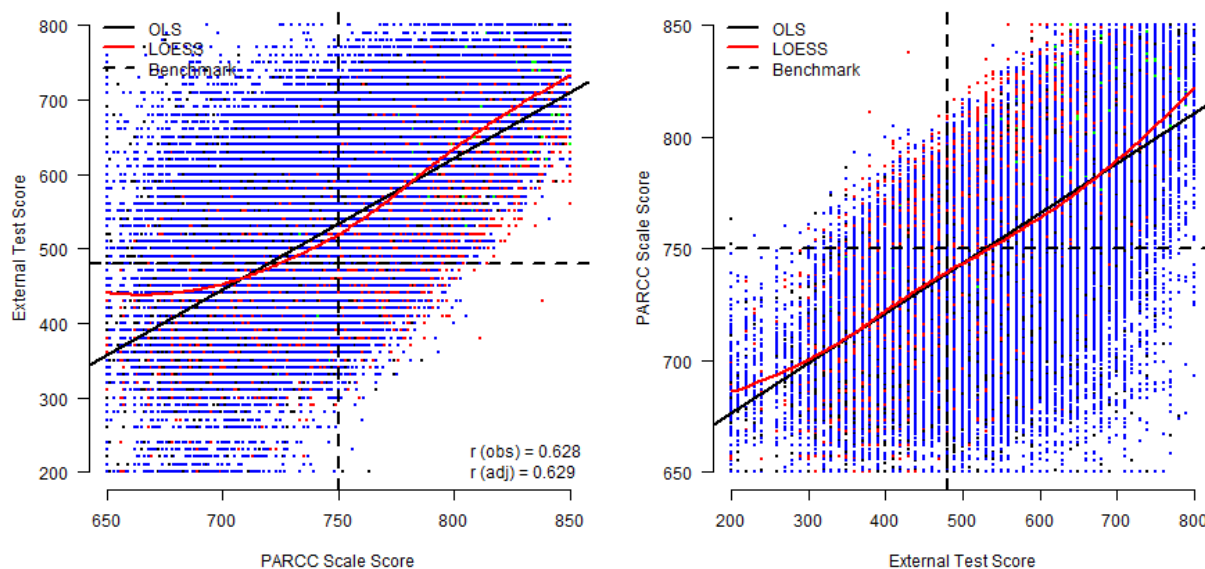| Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Grade 11 ELA/L	750	82,201	533.5	519.9	.750	.692
SAT EBRW	480	82,201	739.1	739.9	.365	.401

Figure 3.23. OLS and LOESS regression for Grade 11 ELA/L and SAT EBRW.
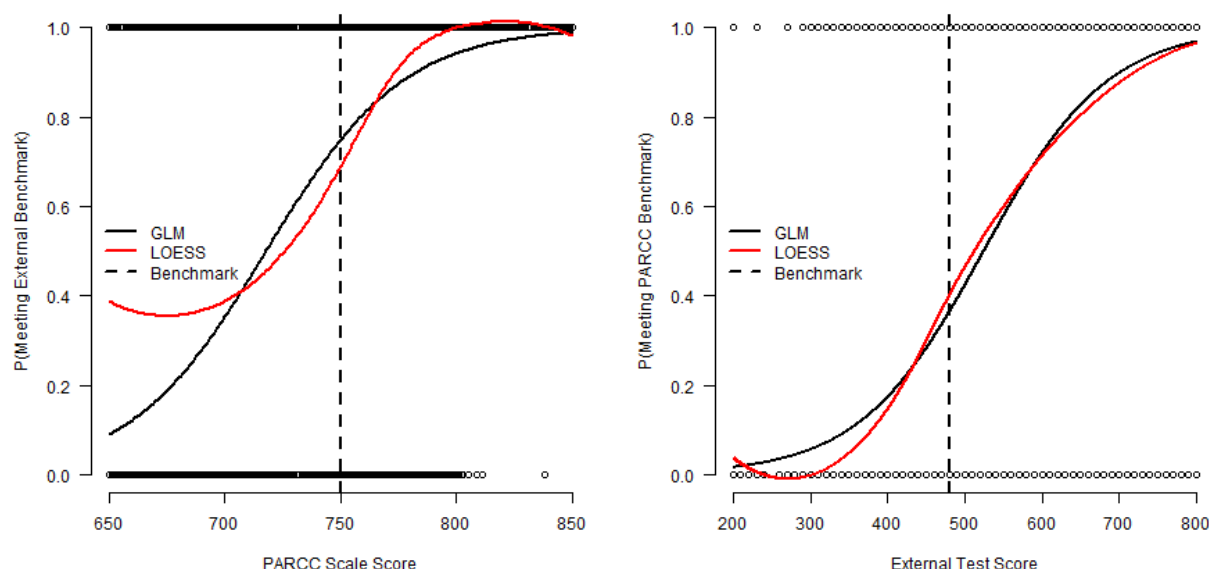
Figure 3.24. Logistic regression for Grade 11 ELA/L and SAT EBRW.

Because the sample of test takers was split between students who took the tests concurrently and those who took PARCC Grade 11 ELA/L first, the analyses were re-run separately for those two groups. In both cases (Tables 3.59 and 3.60), the results were quite consistent with those from the full sample. That is, students meeting the PARCC benchmark were expected to exceed the PSAT/NMSQT benchmark, and students meeting the PSAT/NMSQT benchmark were expected to score below the PARCC benchmark.

Table 3.59
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Grade 11 ELA/L and SAT Evidence-Based Reading and Writing (Only Students Who Took PARCC and SAT Concurrently)

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Grade 11 ELA/L	750	44,116	532.9	517.7	.743	.675
SAT EBRW	480	44,116	740.5	741.6	.388	.429

Table 3.60
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Grade 11 ELA/L and SAT Evidence-Based Reading and Writing (Only Students Who PARCC Before SAT)

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Grade 11 ELA/L	750	33,179	528.9	518.1	.746	.697
SAT EBRW	480	33,179	737.4	737.6	.332	.362

3.6.4 Grade 11 ELA/L and ACT English

The sample of students that took PARCC Grade 11 ELA/L and the ACT English test was more able, on average, than the PARCC Grade 11 ELA/L test taking population (Table 3.61). This sample had more Hispanic students, but fewer Black students, White students, and students with disabilities than the population. The sample was split between students who took the tests concurrently as eleventh graders (54.8%) and students who took the PARCC assessment as eleventh graders and the SAT in the fall of twelfth grade (28.0%; Table 3.62).

Table 3.61

Demographic Comparison for Sample Taking PARCC Grade 11 ELA/L and ACT English

Variable	PARCC Population	Sample	Difference	Effect Size
N	304,214	6,635		
% Female	51.6	41.2	-10.4	-0.21
% American Indian/Alaskan Native	1.9	9.5	7.6	0.35
% Asian	5.2	2.8	-2.5	-0.13
% Black/African American	14.3	1.6	-12.7	-0.52
% Hispanic	27.0	50.4	23.3	0.49
% Native Hawaiian/Other Pacific Islander	0.2	0.1	-0.1	-0.02
% White	47.7	34.9	-12.7	-0.26
% English Language Learners	3.9	2.8	-1.1	-0.06
% Economically Disadvantaged	40.0	45.7	5.7	0.12
% Students with Disability	12.3	3.2	-9.0	-0.35
% PARCC Level 4 or Level 5	38.8	67.9	29.1	0.59
Scale Score Mean	738.2	763.9	25.7	0.67
Scale Score SD	38.7	31.4		

Table 3.62

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.0	0.2	0.3	54.8	2.4
External First	0.0	0.2	0.1	10.7	1.6
PARCC First	0.0	0.0	0.3	28.0	1.4

A lower percentage of students met the ACT English benchmark than met the PARCC benchmark (53.9% vs. 67.9%; Table 3.63). Students who earned a score of 750 on PARCC Grade 11 ELA/L were expected to score 17.1 on ACT English, which is just below the ACT benchmark of 18 (Table 3.63 and Figure 3.25). Students who just met the ACT benchmark were expected to score 760.3, which is above the PARCC benchmark score of 750. Logistic regression indicated that students who scored 750 on PARCC Grade 11 ELA/L had a 0.361 probability of meeting the

ACT benchmark, and students who just met the ACT benchmark had a 0.725 probability of meeting the PARCC benchmark (Table 3.64 and Figure 3.26). The adjusted correlation between scores was 0.788.

Nearly all of the data for this analysis was from State D. Students may have performed unexpectedly well on the PARCC assessment considering their ACT performance (e.g., if meeting certain level of PARCC performance was required for high school graduation in State D but not other states). That could potentially explain why, unlike the PSAT and SAT, more students met the PARCC benchmark than met the external assessment benchmark.

Table 3.63
Cross-Tabulation Percentages for Sample Taking PARCC Grade 11 ELA/L and ACT English

	Below Level 4	Level 4 or Level 5	
Below Benchmark	26.4	19.7	46.1
Met Benchmark	5.7	48.2	53.9
	32.1	67.9	

Table 3.64
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Grade 11 ELA/L and ACT English

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Grade 11 ELA/L	750	6,635	17.1	16.3	.361	.315
ACT English	18	6,635	760.3	761.2	.725	.754

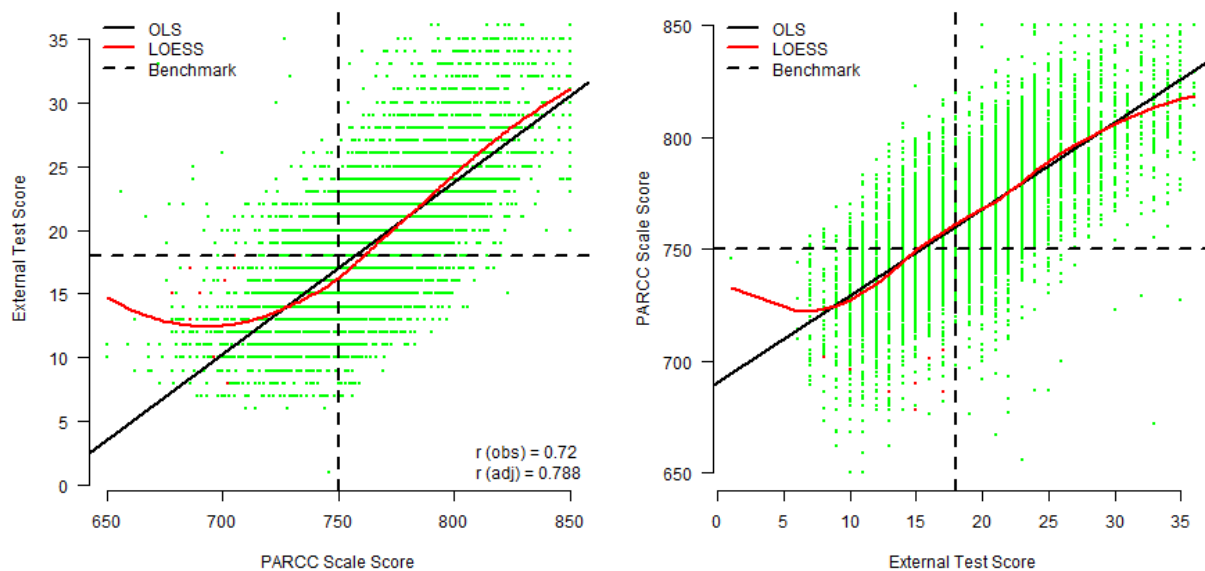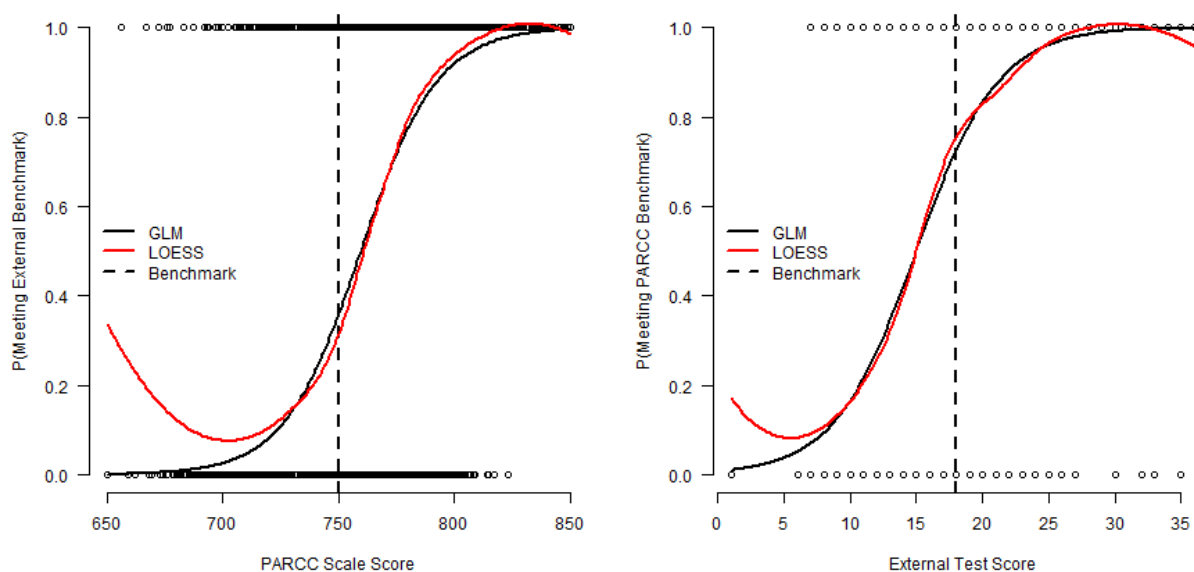*Figure 3.25*. OLS and LOESS regression for Grade 11 ELA/L and ACT English.

Figure 3.26. Logistic regression for Grade 11 ELA/L and ACT English.

Because the sample of test takers was split between students who took the tests concurrently and those who took PARCC Grade 11 ELA/L first, the analyses were re-run separately for those two groups. In both cases (Tables 3.65 and 3.66), the results were quite consistent with those

from the full sample. That is, students meeting the PARCC benchmark were expected to score just below the ACT benchmark, and students meeting the PSAT/NMSQT benchmark were expected to exceed the PARCC benchmark.

Table 3.65
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Grade 11 ELA/L and ACT English (Only Students Who Took PARCC and ACT Concurrently)

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Grade 11 ELA/L	750	3,824	17.0	16.2	.343	.302
ACT English	18	3,824	762.1	762.1	.760	.789

Table 3.66
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Grade 11 ELA/L and ACT English (Only Students Who PARCC Before ACT)

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Grade 11 ELA/L	750	1,975	17.2	16.5	.384	.323
ACT English	18	1,975	756.9	758.2	.658	.678

3.6.5 Grade 11 ELA/L and ACT Reading

There was a high ability sample of students that took PARCC Grade 11 ELA/L and the ACT Reading test (Table 3.67). Compared to the population of PARCC Grade 11 ELA/L test takers, this sample had more Hispanic students, but fewer Black students, White students, and students with disabilities than the population. The sample was split between students who took the tests concurrently as eleventh graders (54.8%) and students who took the PARCC assessment as eleventh graders and the SAT in the fall of twelfth grade (27.9%; Table 3.68).

Table 3.67

Demographic Comparison for Sample Taking PARCC Grade 11 ELA/L and ACT Reading

Variable	PARCC Population	Sample	Difference	Effect Size
N	304,214	6,781		
% Female	51.6	41.1	-10.5	-0.21
% American Indian/Alaskan Native	1.9	9.8	7.9	0.36
% Asian	5.2	2.7	-2.5	-0.13
% Black/African American	14.3	1.6	-12.7	-0.52
% Hispanic	27.0	50.1	23.1	0.48
% Native Hawaiian/Other Pacific Islander	0.2	0.1	-0.1	-0.02
% White	47.7	35.0	-12.7	-0.26
% English Language Learners	3.9	2.8	-1.1	-0.06
% Economically Disadvantaged	40.0	45.6	5.7	0.12
% Students with Disability	12.3	3.2	-9.0	-0.35
% PARCC Level 4 or Level 5	38.8	68.1	29.3	0.60
Scale Score Mean	738.2	764.2	26.1	0.67
Scale Score SD	38.7	31.6		

Table 3.68

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.0	0.2	0.3	54.8	2.2
External First	0.0	0.2	0.1	10.9	1.5
PARCC First	0.0	0.0	0.3	27.9	1.5

A lower percentage of students met the ACT Reading benchmark than met the PARCC benchmark (41.4% vs. 68.1%; Table 3.69). Students who earned a score of 750 on PARCC Grade 11 ELA/L were expected to score 18.9 on ACT Reading, which is below the ACT benchmark of 22 (Table 3.69 and Figure 3.27). Students who just met the ACT benchmark were expected to score 768.9, which is above the PARCC benchmark score of 750. Logistic regression indicated that students who scored 750 on PARCC Grade 11 ELA/L had a 0.198 probability of meeting the ACT benchmark, and students who just met the ACT benchmark had a 0.833 probability of meeting the PARCC benchmark (Table 3.70 and Figure 3.28). The adjusted correlation between scores was 0.773.

Nearly all of the data for this analysis was from State D. Students may have performed unexpectedly well on the PARCC assessment considering their ACT performance (e.g., if meeting certain level of PARCC performance was required for high school graduation in State D

but not other states) . That could potentially explain why, unlike the PSAT and SAT, more students met the PARCC benchmark than met the external assessment benchmark.

Table 3.69
Cross-Tabulation Percentages for Sample Taking PARCC Grade 11 ELA/L and ACT Reading

	Below Level 4	Level 4 or Level 5	
Below Benchmark	28.9	29.7	58.6
Met Benchmark	3.0	38.4	41.4
	31.9	68.1	

Table 3.70
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Grade 11 ELA/L and ACT Reading

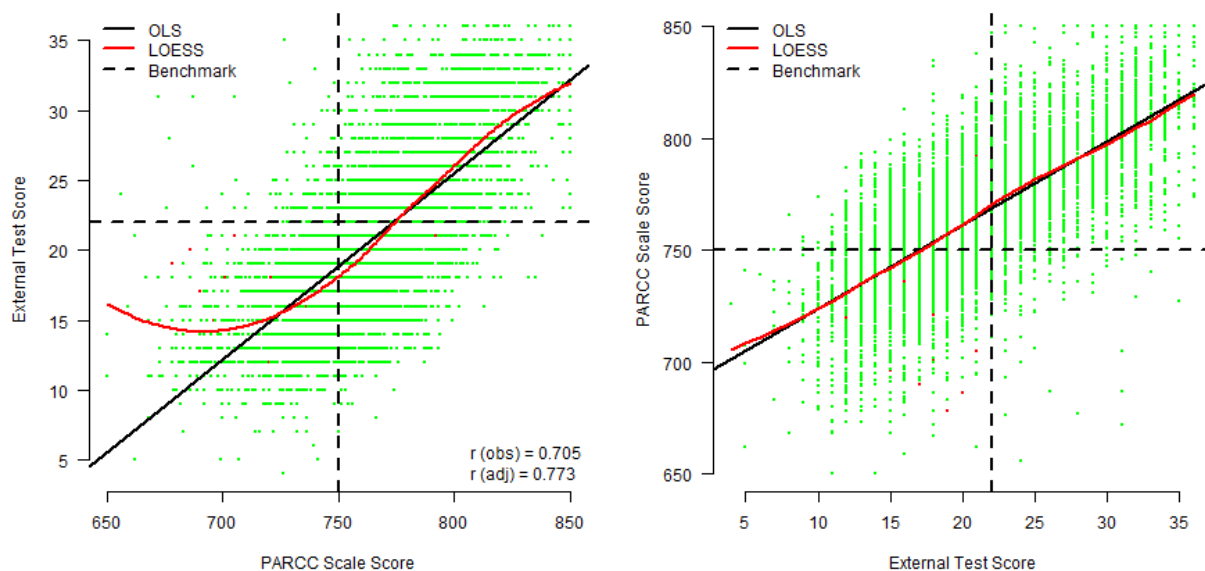| Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Grade 11 ELA/L	750	6,781	18.9	18.2	.198	.189
ACT Reading	22	6,781	768.9	770.4	.833	.842

Figure 3.27. OLS and LOESS regression for Grade 11 ELA/L and ACT Reading.
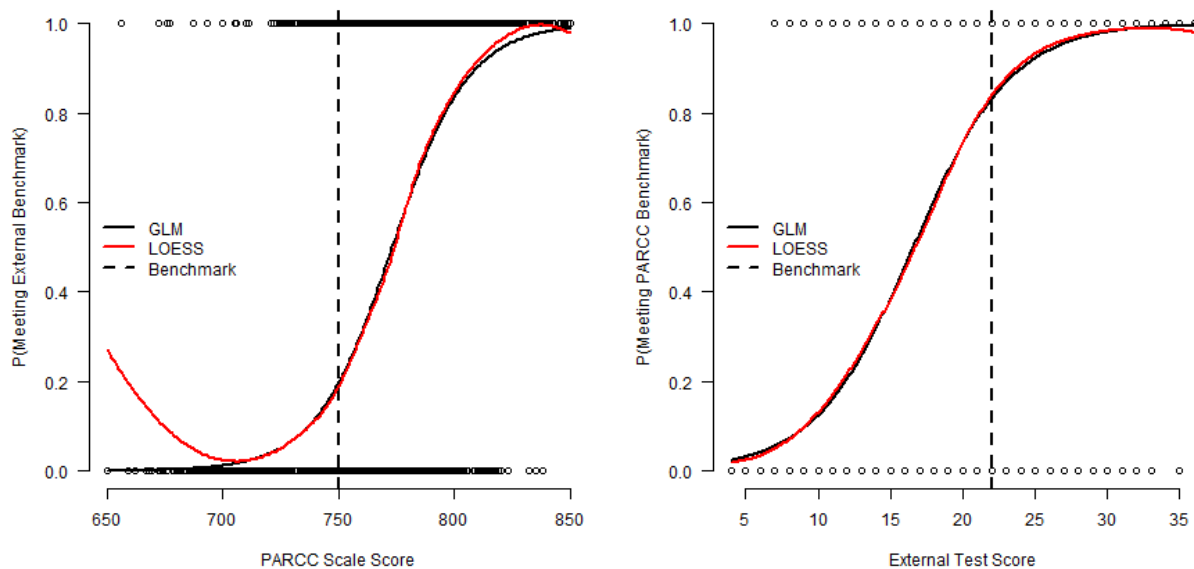
Figure 3.28. Logistic regression for Grade 11 ELA/L and ACT Reading.

Because the sample of test takers was split between students who took the tests concurrently and those who took PARCC Grade 11 ELA/L first, the analyses were re-run separately for those two groups. In both cases (Tables 3.71 and 3.72), the results were quite consistent with those from the full sample. That is, students meeting the PARCC benchmark were expected to score below the ACT benchmark, and students meeting the PSAT/NMSQT benchmark were expected to exceed the PARCC benchmark.

Table 3.71
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Grade 11 ELA/L and ACT Reading (Only Students Who Took PARCC and ACT Concurrently)

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Grade 11 ELA/L	750	3,904	18.7	18.0	.185	.187
ACT Reading	22	3,904	771.0	772.2	.858	.862

Table 3.72
Expected Scores and Probabilities of Meeting Benchmarks for Sample Taking PARCC Grade 11 ELA/L and ACT Reading (Only Students Who PARCC Before ACT)

Test	Benchmark	N	Exp. Score	Exp. Score (LOESS)	Prob.	Prob. (LOESS)
Grade 11 ELA/L	750	2,015	19.0	18.4	.209	.198
ACT Reading	22	2,015	764.9	766.3	.782	.793

4. Discussion

This longitudinal study of external validity of the PARCC performance levels aims to evaluate the validity of the PARCC Level 4 performance level as an indicator of being college ready or "on track" to being college ready. In the first phase of this study, PARCC high school assessment data from the 2014–2015 and 2015–2016 academic years were used to examine the associations between the PARCC performance levels and college readiness benchmarks established by the College Board and ACT. The findings from Phase 1 suggest that the PARCC benchmark score (750) is more difficult to achieve than the benchmark scores for external tests linked to college readiness. The majority of comparisons between performance on PARCC and external measures indicated that a student achieving the PARCC benchmark had a high probability of meeting the benchmark on the external tests. However, students achieving the benchmark on the external test were not as likely to attain PARCC Level 4 or higher.

The second phase of this study (to occur in 2018) will use longitudinal data set including academic performance in entry-level college courses for students who took PARCC assessments during high school. This second phase may lend additional support for the validity of the PARCC Level 4 performance level as a predictor of postsecondary academic success.

Math Results Summary

Given the available data, there were six main bivariate analyses conducted on math assessments to explore the validity of the PARCC Level 4 performance level as an indicator of college readiness: Algebra 1 and PSAT10 Math; Geometry and PSAT10 Math; Algebra II and PSAT10 Math; Algebra II and PSAT/NMSQT Math; Algebra II and SAT Math; and, Algebra II and ACT Math. Table 4.1 summarizes results for these assessment combinations. The nonparametric (LOESS) procedure results were similar to the parametric results, so only the parametric test results are reported in the summary table for clarity.

The classification agreement (meeting the benchmark on both tests or not meeting the benchmark on both tests) ranged from 62.5% to 86.5%, suggesting that the two tests sorted examinees similarly in the majority of cases. With adjusted correlations between scores ranging from .612 to .767, there was a moderate positive relationship between scores on the external and PARCC tests.

The overall trend across all but one of the analyses showed that students who met the benchmark on a PARCC mathematics assessment were likely to meet or exceed the benchmark on an external test (probabilities ranging from .509 to .886). However, students who met the benchmark on the external test had low probabilities of meeting the benchmark on PARCC

(probabilities ranging from .097 to .310). These findings are consistent with the notion that the PARCC benchmark indicates a higher level of achievement than the external test benchmarks.

Algebra I vs. PSAT10 Math was the exception to the general trend in results. A student meeting the benchmark on the Algebra I had a 28.5% chance of meeting or exceeding the benchmark on the external test. Conversely, a student meeting the benchmark on the PSAT10 had a 40.3% chance of attaining at least Level 4 on Algebra I. This sample had lower average ability (-0.26 SD) than the population of PARCC Algebra I test takers. Scores on both the external test and the PARCC assessment were relatively low for these examinees; less than 18% met the benchmark on either assessment. Unlike the other math assessment comparisons, the majority of these examinees (60%) took the PARCC assessment first, which may explain the difference in results compared to the other mathematics test analyses.

Table 4.1
Math results summary.

Test	Benchmark	N	Pred. Score	Prob.	Adj. corr.	Met benchmark	Met both benchmarks	Classification agreement	Sample mean (ES)	Testing order (%)
Algebra I	750	42,637	437.8	.285	.612	18.3%	9.4%	84.6%	725.0 (-.26)	PARCC first[1] (60.3)
PSAT10 Math	480		740.9	.403		16.0%				
Geometry	750	15,779	486.8	.665	.663	18.8%	15.8%	79.2%	727.7 (-.09)	External first[1] (88.4)
PSAT10 Math	480		735.5	.203		33.7%				
Algebra II	750	51,259	549.2	.886	.731	40.0%	38.6%	62.5%	739.3 (.51)	External first[1] (78.3)
PSAT10 Math	480		726.1	.151		74.8%				
Algebra II	750	62,775	553.8	.760	.718	19.6%	18.1%	72.5%	718.7 (-.04)	External first[1] (60.0)
PSAT/NMSQT Math	510		722.8	.113		44.1%				
Algebra II	750	59,154	563.8	.697	.711	13.0%	11.7%	77.2%	711.7 (-.22)	Concurr. (50.2)
SAT Math	530		720.8	.097		33.2%				PARCC 1st (45.3)
Algebra II	750	6,627	21.3	.509	.767	17.8%	14.1%	86.5%	721.1 (.03)	Concurr. (47.6)
ACT Math	22		739.4	.310		23.8%				PARCC 1st (39.4)

[1]Concurrent testing < 3%.

Test = PARCC or external (SAT, PSAT, or ACT)
Benchmark = college readiness benchmark score
N = sample size
Pred Score = predicted score on other assessment for students achieving benchmark
Prob = probability of a student meeting the benchmark on the other assessment
Adj corr = adjusted correlation between assessments
Met benchmark = percentage of examinees who met the benchmark
Met both benchmarks = percentage of examinees who met the benchmark on both assessments
Classification agreement = percentage of examinees whose results matched (they either met/did not meet benchmarks on both tests)
Sample mean (ES) = mean PARCC scale score of the sample and the difference in standard deviation units between sample and population (effect size)

ELA/L Results Summary

Eight bivariate analyses on ELA/L assessments were conducted to explore the validity of the PARCC Level 4 performance level as an indicator of college readiness: grade 9 and PSAT10 EBRW; grade 10 and PSAT10 EBRW; grade 10 and PSAT/NMSQT EBRW; grade 10 and SAT EBRW; grade 11 and PSAT/NMSQT EBRW; grade 11 and SAT EBRW; grade 11 and ACT English; and grade 11 and ACT reading. Table 4.2 summarizes results for these assessment combinations. The nonparametric (LOESS) procedure results were similar to the parametric results, so only the parametric test results are reported in the summary table for clarity.

The classification agreement (meeting the benchmark on both tests or not meeting the benchmark on both tests) ranged from 67.3% to 79.7% on both assessments, suggesting that the two tests sorted examinees similarly in the majority of cases. The adjusted correlation between scores ranged from .629 to .788, indicating a moderate positive relationship between scores on the external and PARCC tests. These results are in line with the math results discussed in the previous section.

With the exception of both ACT comparisons, the trend in ELA/L results indicated that students meeting the PARCC ELA/L benchmarks had between a 66.7% and 82.5% chance of meeting the benchmark on the external assessment. However, a student taking the external test had between a 32.6% and 51.3% chance of meeting the PARCC benchmark.

The results for the ACT and grade 11 ELA/L test did not follow the general trend. The ACT sample was higher ability than the PARCC population by an average of 0.67 standard deviations, and the comparison group came primarily from State D. These students may have performed unexpectedly poorly on the ACT considering their PARCC performance. This could potentially explain why, unlike the PSAT and SAT groups, more students met the PARCC benchmark than met the external assessment benchmark.

Table 4.2
ELA results summary.

Test	Benchmark	N	Pred. Score	Adj. Corr.	Prob.	Met benchmark	Met both benchmarks	Classification agreement	Sample mean (ES)	Testing order (%)
Grade 9 ELA/L	750	8,836	471.1	.652	.684	71.1%	57.3%	78.3%	762.5	PARCC first
PSAT10 EBRW	430		747.9		.513	65.3%			(.67)	(91.5)
Grade 10 ELA/L	750	122,251	468	.733	.692	50.3%	44.9%	78.0%	749.3	External first
PSAT10 EBRW	430		737.7		.357	61.5%			(.29)	(97.5)
Grade 10 ELA/L	750	47,170	511.2	.723	.775	53.2%	49.3%	75.0%	752.5	PARCC first
PSAT/NMSQT EBRW	460		736		.326	70.4%			(.36)	(97.4)
Grade 10 ELA/L	750	10,195	546.2	.704	.825	65.1%	62.1%	79.7%	765.0	PARCC 1st (65.2)
SAT EBRW	480		738.5		.385	79.5%			(.63)	Concurr. (29.3)
Grade 11 ELA/L	750	64,823	497.6	.635	.667	50.9%	43.1%	73.0%	750.2	External first
PSAT/NMSQT EBRW	460		741.2		.395	62.3%			(.31)	(93.1)
Grade 11 ELA/L	750	82,201	533.5	.629	.750	52.7%	47.1%	70.9%	751.7	Concurr. (53.7)
SAT EBRW	480		739.1		.365	70.5%			(.35)	PARCC 1st (40.3)
Grade 11 ELA/L	750	6,635	17.1	.788	.361	67.9%	48.2%	74.6%	763.9	Concurr. (57.7)
ACT English	18		760.3		.725	53.9%			(.67)	PARCC 1st (29.7)
Grade 11 ELA/L	750	6,781	18.9	.773	.198	68.1%	38.4%	67.3%	764.2	Concurr. (57.5)
ACT Reading	22		768.9		.833	41.4%			(.67)	PARCC 1st (29.7)

Test = PARCC or external (SAT, PSAT, or ACT)
Benchmark = college readiness benchmark score
N = sample size
Pred Score = predicted score on other assessment for students achieving benchmark
Prob = probability of a student meeting the benchmark on the other assessment; Adj corr = adjusted correlation between assessments
Met benchmark = percentage of examinees who met the benchmark
Met both benchmarks = percentage of examinees who met the benchmark on both assessments
Classification agreement = percentage of examinees whose results matched (they either met/did not meet benchmarks on both tests)
Sample mean (ES) = mean PARCC scale score of the sample and the difference in standard deviation units between sample and population (effect size)

Conclusion and Implications

This study addressed the validity of PARCC Level 4 as an indicator of being college ready by examining associations with College Board and ACT benchmarks for college readiness. This study comprised a series of comparisons between PARCC Level 4 and college readiness benchmarks on external tests, summarized in Table 4.1 and 4.2. Results indicated that a student meeting the benchmark on the PARCC test had a high probability of making the benchmark on the external test, but the converse did not hold for students meeting the benchmark on the external test. The classification agreement between tests averaged 76% (SD=6.3%). The adjusted correlation between tests averaged .70 (SD = .06). For comparison, the correlation between NAEP and SAT is .91 and .74 for math and critical reading, respectively (Camara, 2013). A correlation of .8 or greater between assessments is considered high when using one score to predict the outcome on the other (Dorans, 1999).

Interpretation of results from this study are limited by the available study data. There was variability among the available samples for this study. Some samples were high-ability compared to the population of PARCC examinees. For example, in terms of mean ability, the Algebra II sample was 0.5 standard deviations above the population, and the grade 9 ELA/L sample was 0.67 standard deviations above the population. All of the ELA/L test takers in the samples had higher average ability compared to the PARCC population. Some samples of math examinees were lower in ability than the PARCC population. The study samples also differed from the populations in terms of some demographic variables. Typically at least one subgroup (e.g., black, Hispanic, English language learner, etc.) was at least one quarter of a standard deviation above or below the PARCC population.

Moreover, the order in which the tests were administered was not consistent, and this must be considered when interpreting results. Some comparisons involved students who took one test at the beginning of a grade (in the fall) and the other test at the end of the grade (in the spring). Learning would be expected over the course of that academic year. For example, students who took Algebra II before the SAT were expected to score higher on the SAT than students who took the tests concurrently (section 3.34). This result may be explained by the fact that students are likely to perform better on the SAT in grade 12 than in grade 11.

Student motivation is another factor to consider when interpreting results from this study. Motivation is known to be associated with assessment performance, and students tend to perform better when stakes are attached to performance (Wise & DeMars, 2005). For example, students may be motivated to perform well on tests like the SAT and ACT because scores are considered in college admissions decisions. In contrast, there are not always stakes attached to PARCC performance. In some states, students may need to demonstrate a specific level of performance on certain PARCC assessments to graduate from high school. Students in those

states may perform better on PARCC than similar students in states without a high school graduation requirement attached to PARCC performance. The results of this study for the ACT English and Reading assessments potentially illustrate this idea. In all, factors such as sample representation, assessment timing, and motivation warrant caution when interpreting results and attempting to generalize from this study to the broader PARCC population.

Any comparisons of performance on different assessments depend on the extent to which assessments measure similar content. PARCC, College Board, and ACT assessments do not measure the same body of content. However, convergent performance supports the validity of PARCC benchmarks. Research conducted by the Massachusetts Executive Office of Education on the relationship between Massachusetts Comprehensive Assessment System (MCAS) and PARCC found that both MCAS and PARCC predict college readiness. In mathematics, meeting the PARCC benchmark predicted a higher level of performance in college than meeting the MCAS benchmark, whereas in ELA/L, the two tests were comparable in predicting college performance. The MCAS study found that the majority of test takers meeting the benchmark on PARCC (93% in ELA/L and 84% in mathematics) earned at least a C in their first-year college courses in these subjects. In addition, PARCC's standard was a better predictor than MCAS in identifying students who will earn a B or higher GPA in college (Nichols-Barrer, Place, Dillon, & Gill, 2015). To date, this is the only study linking PARCC performance to actual college outcomes.

The Maryland Assessment Research Center conducted a study using data from grade 12 students' College Board and ACT scores matched with PARCC Algebra I, Algebra II, and grade 10 ELA/L. Some, but not all, of the results aligned with the findings from this PARCC validation study. In the Maryland study, students achieving the PARCC level 4 benchmark on Algebra II had above-benchmark scores on SAT and ACT math. However, students achieving PARCC level 4 benchmark on Grade 10 ELA/L did not meet the SAT Reading or Writing benchmarks. The Maryland study sample size was small, the students scored lower on the ability scale, and the timing of the tests was not clear. On average, the results from the Maryland study indicated that students with higher scores on PARCC tended to score higher on external tests, and that PARCC and PSAT as predictors together explained approximately 30–60% of the variance in College Board and ACT scores (Liao, Li, Jiao, & Lissitz, 2015).

Results from this first phase of this longitudinal study have several implications. Consistent with prior research, meeting the PARCC benchmark is likely indicative of academic readiness for college (or being "on track"). Students who meet the PARCC benchmark may have a greater than .75 probability of earning a C or higher in first-year college courses. The validity of these preliminary findings cannot be confirmed until phase 2 of this longitudinal study has been conducted. In phase 2, PARCC scores will be correlated with examinee academic performance

in entry-level college courses. With this data, performance levels on PARCC assessments can be directly associated with postsecondary academic outcomes.

References

Allen, J. (2013). Updating the ACT college readiness benchmarks. Iowa City, IA: ACT. Retrieved from http://files.eric.ed.gov/fulltext/ED546851.pdf

Camara, W. (2013). Defining and measuring college and career readiness: A validation framework. Educational Measurement: Issues and Practice, 32(4), 16-27.

CCSSO. (2014). Criteria for procuring and evaluating high-quality assessments. Retrieved January 27, 2017 from http://www.ccsso.org/Documents/2014/CCSSO%20Criteria%20for%20High%20Quality%20Assessments%2003242014.pdf

College Board. (2015). PSAT/NMSQT: Understanding scores 2015. New York, NY: College Board. Retrieved from https://collegereadiness.collegeboard.org/pdf/2015-psat-nmsqt-understanding-scores.pdf

Dorans, N. J. (1999). Correspondences between ACT™ and SAT® I scores. ETS Research Report Series, 1999(1).

Liao, D., Li, C., Jiao, H., & Lissitz, R. W. (2015). Investigating the relationship between the PARCC test scores and the college admission tests: SAT/ACT/PSAT (MARC Technical Report). College Park, MD: University of Maryland.

Nichols-Barrer, I., Place, K., Dillon, E., & Gill, B. (2015). Predictive Validity of MCAS and PARCC: Comparing 10th Grade MCAS Tests to PARCC Integrated Math II, Algebra II, and 10th Grade English Language Arts Tests (No. a2d9543914654aa5b012e4a6d2dae060). Mathematica Policy Research.

Schultz, S.R., Michaels, H.R., Dvorak, R.N., & Wiley, C.R. (2016). Evaluating the content and quality of next generation high school assessments. Retrieved January 27, 2017 from https://www.humrro.org/corpsite/sites/default/files/HQAP_HumRRO_High_School_Study_Final%20Report.pdf

Shaw, E.J., Marini, J.P., Beard, J., Shmueli, D., Young, L., & Ng, H. (2016). The redesigned SAT® pilot predictive validity study: A first look New York, NY: College Board. Retrieved from https://collegereadiness.collegeboard.org/pdf/redesigned-sat-pilot-predictive-validity-study-first-look.pdf

Wise, S.L., & DeMars, C.E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment*, *10*(1), 1-17. doi: 10.1207/s15326977ea1001_1

Appendix A

Table A.1. Full OLS results.

Test	Benchmark	N	Exp. Score (SE)	Prob.	Adj. Corr.	Intercept (SE)	Slope (SE)
Algebra I	750	42,637	437.8 (.41)	.285	.612	435.3 (7.93)	1.5 (.01)
PSAT10 Math	480		740.9 (.17)	.403		742.1 (.62)	.2 (<.01)
Geometry	750	15,779	486.8 (.80)	.665	.663	-1368.5 (17.63)	2.5 (.02)
PSAT10 Math	480		735.5 (.17)	.203		658.3 (.70)	.2 (<.01)
Algebra II	750	51,259	549.2 (.31)	.886	.731	-857.8 (6.09)	1.9 (.01)
PSAT10 Math	480		726.1 (.13)	.151		597.3 (.63)	.3 (<.01)
Algebra II	750	62,775	553.8 (.38)	.760	.718	-894.5 (5.79)	1.9 (.01)
PSAT Math	510		722.8 (.10)	.113		596.6 (.52)	0.2 (<.01)
Algebra II	750	59,154	563.8 (.46)	.697	.711	-898.2 (6.48)	1.9 (.01)
SAT Math	530		720.8 (.11)	.097		602.1 (.52)	0.2 (<.01)
Algebra II	750	6,627	21.3 (.05)	.509	.767	-47.3 (.81)	0.1 (<.01)
ACT Math	22		739.4 (.36)	.310		619.7 (1.28)	5.4 (.07)
Grade 9 ELA/L	750	8,836	471.1 (1.0)	.684	.652	-1126.7 (18.81)	2.1 (.02)
PSAT10 EBRW	430		747.9 (.35)	.513		655.4 (1.28)	0.2 (<.01)
Grade 10 ELA/L	750	122,251	468.0 (.21)	.692	.733	-797.0 (3.40)	1.7 (<.01)
PSAT10 EBRW	430		737.7 (.09)	.357		602.0 (.40)	0.3 (<.01)
Grade 10 ELA/L	750	47,170	511.2 (.34)	.775	.723	-756.8 (5.83)	1.7 (.01)
PSAT EBRW	460		736.0 (.16)	.326		599.1 (.72)	0.3 (<.01)
Grade 10 ELA/L	750	10,195	546.2 (.86)	.825	.704	-801.3 (13.33)	1.8 (.02)
SAT EBRW	480		738.5 (.42)	.385		601.8 (1.61)	0.3 (<.01)
Grade 11 ELA/L	750	64,823	497.6 (.32)	.667	.635	-812.6 (6.12)	1.7 (.01)
PSAT/NMSQT EBRW	460		741.2 (.13)	.395		631.8 (.56)	0.2 (<.01)
Grade 11 ELA/L	750	82,201	533.5 (.29)	.750	.629	-786.7 (5.73)	1.8 (.01)
SAT EBRWS	480		739.1 (.12)	.365		631.6 (.53)	0.2 (<.01)
Grade 11 ELA/L	750	6,635	17.1 (.05)	.361	.788	-83.7 (1.22)	0.1 (<.01)
ACT English	18		760.3 (.27)	.725		690.8 (.91)	3.9 (.05)
Grade 11 ELA/L	750	6,781	18.9 (.06)	.198	.773	-81.0 (1.24)	0.1 (<.01)
ACT Reading	22		768.9 (.28)	.833		686.8 (.99)	3.7 (.05)

Appendix B

Comparisons that were considered but for which the data did not support an analysis.

B.1 Algebra I and PSAT/NMSQT Math

There was a fairly small sample of students who took PARCC Algebra I and the PSAT/NMSQT within six months of each other. The sample was of lower average ability (Table B.1) and took Algebra I as tenth or eleventh graders (Table B.2), which is later than the majority of students. For these reasons, results for these data are not reported.

Table B.1

Demographic Comparison for Sample Taking PARCC Algebra I and PSAT/NMSQT Math

Variable	PARCC Population	Sample	Difference	Effect Size
N	624,008	2,437		
% Female	51.7	56.5	4.8	0.10
% American Indian/Alaskan Native	1.2	0.3	-0.9	-0.11
% Asian	6.0	4.3	-1.6	-0.08
% Black/African American	18.3	30.5	12.2	0.29
% Hispanic	28.3	21.9	-6.4	-0.15
% Native Hawaiian/Other Pacific Islander	0.2	0.2	0.0	0.00
% White	42.2	40.2	-2.0	-0.04
% English Language Learners	7.3	13.6	6.3	0.21
% Economically Disadvantaged	43.2	49.0	5.8	0.12
% Students with Disability	12.5	26.8	14.4	0.37
% PARCC Level 4 or Level 5	31.3	9.8	-21.4	-0.55
Scale Score Mean	733.6	716.1	-17.5	-0.52
Scale Score SD	33.7	25.7		

Table B.2

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.0	0.0	0.0	1.4	0.1
External First	0.0	0.0	0.0	38.7	1.4
PARCC First	0.0	0.0	53.2	5.3	0.0

B.2 Algebra I and SAT Math

There was a small sample of students who took PARCC Algebra I and the SAT within six months of each other. The sample was of lower average ability (Table B.3) and took Algebra I mainly as eleventh graders (Table B.4), which is later than the majority of students. For these reasons, results for these data are not reported.

Table B.3

Demographic Comparison for Sample Taking PARCC Algebra I and SAT Math

Variable	PARCC Population	Sample	Difference	Effect Size
N	624,008	1,080		
% Female	51.7	51.0	-0.7	-0.01
% American Indian/Alaskan Native	1.2	0.4	-0.8	-0.10
% Asian	6.0	5.6	-0.4	-0.02
% Black/African American	18.3	45.9	27.7	0.61
% Hispanic	28.3	21.8	-6.5	-0.15
% Native Hawaiian/Other Pacific Islander	0.2	0.1	-0.1	-0.02
% White	42.2	22.1	-20.1	-0.44
% English Language Learners	7.3	12.3	5.0	0.17
% Economically Disadvantaged	43.2	47.6	4.4	0.09
% Students with Disability	12.5	13.4	0.9	0.03
% PARCC Level 4 or Level 5	31.3	15.2	-16.1	-0.39
Scale Score Mean	733.6	720.6	-13.0	-0.39
Scale Score SD	33.7	29.3		

Table B.4

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.0	1.9	4.7	44.3	4.8
External First	0.0	0.3	0.2	1.6	9.4
PARCC First	0.2	1.5	4.8	25.2	1.2

B.3 Algebra I and ACT Math

The data included only 110 students who took PARCC Algebra I and the ACT within six months of each other, many of whom took Algebra I as eleventh or twelfth graders, which is later than the majority of students. Results for these data are not reported.

B.4 Geometry and PSAT/NMSQT Math

There were approximately 5,400 students who took PARCC Geometry and the PSAT/NMSQT within six months of each other. Although the sample size was adequate for analyses, the students had very low average ability (4.6% attaining PARCC Level 4 or 5; Table B.5) and took Geometry primarily as eleventh graders (Table B.6), which is later than the majority of students. For these reasons, results for these data are not reported.

Table B.5

Demographic Comparison for Sample Taking PARCC Geometry and PSAT/NMSQT Math

Variable	PARCC Population	Sample	Difference	Effect Size
N	289,864	5,387		
% Female	51.3	54.2	2.9	0.06
% American Indian/Alaskan Native	1.9	0.8	-1.2	-0.10
% Asian	7.0	4.9	-2.1	-0.09
% Black/African American	12.7	23.1	10.4	0.27
% Hispanic	29.3	32.4	3.0	0.07
% Native Hawaiian/Other Pacific Islander	0.2	0.2	0.0	-0.01
% White	46.4	37.3	-9.1	-0.19
% English Language Learners	5.2	7.4	2.2	0.09
% Economically Disadvantaged	38.4	47.7	9.3	0.19
% Students with Disability	13.0	26.2	13.2	0.34
% PARCC Level 4 or Level 5	23.9	4.6	-19.3	-0.59
Scale Score Mean	730.0	713.8	-16.3	-0.61
Scale Score SD	26.7	19.9		

Table B.6

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.0	0.0	0.0	5.3	0.1
External First	0.0	0.0	0.0	76.2	0.5
PARCC First	0.0	0.0	8.7	9.1	0.0

B.5 Geometry and SAT Math

There was a relatively small sample of students who took PARCC Geometry and the SAT within six months of each other. The sample was of lower average ability (Table B.7) and took Geometry primarily as eleventh graders (Table B.8), which is later than the vast majority of students. For these reasons, results for these data are not reported.

Table B.7

Demographic Comparison for Sample Taking PARCC Geometry and SAT Math

Variable	PARCC Population	Sample	Difference	Effect Size
N	289,864	3,145		
% Female	51.3	48.6	-2.7	-0.05
% American Indian/Alaskan Native	1.9	0.2	-1.7	-0.18
% Asian	7.0	7.3	0.4	0.02
% Black/African American	12.7	25.1	12.4	0.32
% Hispanic	29.3	26.1	-3.2	-0.07
% Native Hawaiian/Other Pacific Islander	0.2	0.1	-0.1	-0.02
% White	46.4	35.6	-10.8	-0.22
% English Language Learners	5.2	6.6	1.5	0.06
% Economically Disadvantaged	38.4	42.5	4.2	0.09
% Students with Disability	13.0	19.6	6.6	0.18
% PARCC Level 4 or Level 5	23.9	6.3	-17.6	-0.52
Scale Score Mean	730.0	715.9	-14.2	-0.53
Scale Score SD	26.7	21.0		

Table B.8

Percentage of Test Takers by PARCC Grade and Assessment Timing

	8	9	10	11	12
Concurrent	0.0	0.1	1.1	47.6	1.1
External First	0.0	0.0	0.1	1.5	2.8
PARCC First	0.0	0.2	1.8	43.5	0.2

B.6 Geometry and ACT Math

The data included only 568 students who took PARCC Geometry and the ACT within six months of each other, many of whom took Geometry as eleventh or twelfth graders, which is later than the vast majority of students. Results for these data will not be reported.

B.7 Grade 9 ELA/L and PSAT/NMSQT Evidence Based Reading and Writing

There were data from only 36 students who took PARCC Grade 9 ELA/L and the PSAT/NMSQT Evidence-Based Reading and Writing test within 6 months of each other. Results for analyses based on these data are not reported.

B.8 Grade 9 ELA/L and SAT Evidence Based Reading and Writing

There were data from only 131 students who took PARCC Grade 9 ELA/L and the SAT Evidence-Based Reading and Writing Test within 6 months of each other. Results for analyses based on these data are not reported.

B.9 Grade 9 ELA/L and ACT English

There were data from only 19 students who took PARCC Grade 9 ELA/L and the ACT English test within 6 months of each other. Results for analyses based on these data are not reported.

B.10 Grade 9 ELA/L and ACT Reading

There were data from only 22 students who took PARCC Grade 9 ELA/L and the ACT Reading test within 6 months of each other. Results for analyses based on these data are not reported.

B.11 Grade 10 ELA/L and ACT English

There were data from only 760 students who took PARCC Grade 10 ELA/L and the ACT English test within 6 months of each other. Results for analyses based on these data will not be reported.

B.12 Grade 10 ELA/L and ACT Reading

There were data from only 763 students who took PARCC Grade 10 ELA/L and the ACT Reading test within 6 months of each other. Results for analyses based on these data will not be reported.

B.13 Grade 11 ELA/L and PSAT10

There were data from only 825 students who took PARCC Grade 11 ELA/L and the PSAT10 Evidence-Based Reading and Writing test within 6 months of each other, and most of those students took Grade 11 ELA/L as tenth graders. Results for analyses based on these data are not reported.